

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 13:50:30 ; Search time 8904.87 Seconds
(without alignments)
2958.109 Million cell updates/sec

Title: US-08-894-356C-1
Perfect score: 1703
Sequence: 1 TCATTATGGAGCAATCCAA.....AAAAAAAAAAAAAAAAAAAA 1703

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_bal.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_om.*
8: gb_ov.*
9: gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_pl1.*
13: gb_pl2.*
14: gb_pl3.*
15: gb_pl4.*
16: em_bal.*
17: em_ba2.*
18: em_fun.*
19: em_htgo_hum.*
20: em_htgo_inv.*
21: em_htgo_rod.*
22: em_htg_hum1.*
23: em_htg_hum2.*
24: em_htg_hum3.*
25: em_htg_hum4.*
26: em_htg_hum5.*
27: em_htg_hum6.*
28: em_htg_hum7.*
29: em_htg_hum8.*
30: em_htg_inv1.*
31: em_htg_inv2.*
32: em_htg_other.*
33: em_htg_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pl.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vi.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vil.*
59: gb_vil2.*
60: gb_htg1.*
61: gb_htg2.*
62: gb_htg3.*
63: gb_htg4.*
64: gb_htg5.*
65: gb_htg6.*
66: gb_htg7.*
67: gb_htg8.*
68: gb_htg9.*
69: gb_htg10.*
70: gb_htg11.*
71: gb_htg12.*
72: gb_htg13.*
73: gb_htg14.*
74: gb_htg15.*
75: gb_htg16.*
76: gb_htg17.*
77: gb_htg18.*
78: gb_htg19.*
79: gb_htg20.*
80: gb_htg21.*
81: gb_htg22.*
82: gb_htg23.*
83: gb_htg24.*
84: gb_htg25.*
85: gb_pr1.*
86: gb_pr2.*
87: gb_pr3.*
88: gb_pr4.*
89: gb_pr5.*
90: gb_pr6.*
91: gb_pr7.*
92: gb_pr8.*
93: gb_pr9.*
94: gb_rcl.*
95: gb_rcl2.*
96: gb_in4.*
97: gb_pr10.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1703	100.0	1703	10	E12753
2	1679	98.6	1679	12	AB010708
3	204	12.0	1508	10	E12757
4	189.2	11.1	1622	10	E12754
5	189.2	11.1	1622	12	AB026494
6	172.2	10.1	1479	10	E12756
7	171.4	10.1	1476	12	AB029340
8	114	6.7	11914	12	AC003027
					E12753 Gentiana
					AB010708 Gentiana
					E12757 Senecio cru
					E12754 Gentiana
					AB026494 Gentiana
					E12756 Perilla oci
					AB029340 Perilla f
					AC003027 Arabidops

9	109.2	6.4	101176	12	AC002560	Genomic s
10	81	4.8	1648	9	AX025514	Sequence
c 11	71.8	4.2	82360	13	AP006066	Arabidops
c 12	71.8	4.2	341064	75	AC074226	Arabidops
c 13	64.2	3.8	66237	12	AB016892	Arabidops
c 14	63	3.7	136047	74	AC069470	Arabidops
c 15	61.8	3.6	85690	12	AB028618	Arabidops
c 16	61.8	3.6	136047	74	AC069470	Arabidops
c 17	51.2	3.0	71623	77	AC087724	Arabidops
c 18	51.2	3.0	85690	12	AB028618	Arabidops
c 19	49	2.9	1141	10	AX083744	Sequence
c 20	45.6	2.7	321003	84	PF0414P3	
c 21	44.8	2.6	192929	60	AC005505	Plasmodiu
c 22	44.4	2.6	269	6	ECA270224	Entodiniu
c 23	44	2.6	46739	13	AP002052	Arabidops
c 24	44	2.6	78379	12	AB006696	Arabidops
c 25	43.8	2.6	191166	80	AL356295	Homo sapi
c 26	43.2	2.5	1141	10	AX083744	Sequence
c 27	42.8	2.5	3314	5	AF163834	Dictyoste
c 28	42.8	2.5	188925	67	AC022050	Homo sapi
c 29	42.4	2.5	129838	75	AC079038	Oryza sat
c 30	42.4	2.5	154425	60	AC008911	Homo sapi
c 31	42.4	2.5	156060	60	AC004153	Plasmodiu
c 32	42.2	2.5	186974	83	AP003071	Homo sapi
c 33	42	2.5	156060	60	AC004153	Plasmodiu
c 34	42	2.5	170921	83	CNS01DXJ	Homo sapi
c 35	42	2.5	202487	83	CNS01DM6	Homo sapi
c 36	42	2.5	202496	91	CNS01DX6	Human chr
c 37	42	2.5	219936	64	AC016299	Homo sapi
c 38	41.8	2.5	2222	9	AR105363	Sequence
c 39	41.8	2.5	2222	97	HUMKEINF	Homo sapiens
c 40	41.8	2.5	197120	79	AL354820	Homo sapi
c 41	41.8	2.5	200543	64	AC016726	Homo sapi
c 42	41.6	2.4	356	14	PSN06DMR	P. sativum P
c 43	41.2	2.4	14955	5	AF000580	Dictyoste
c 44	41.2	2.4	100000	91	AP000096	Homo sapi
c 45	41.2	2.4	100000	91	AP000200	Homo sapi

ALIGNMENTS

RESULT	1
E12753	
LOCUS	E12753 1703 bp DNA PAT 24-JUN-1998
DEFINITION	Gentianatriflora mRNA acyltransferase, complete cds.
ACCESSION	E12753
VERSION	E12753.1 GI:3251585
KEYWORDS	JP 1997070290-A/1.
SOURCE	unidentified.
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 1703)
AUTHORS	Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y., Yonekura, K., Mizutani, M. and Kusumi, T.
TITLE	GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL	Patent: JP 1997070290-A 1 18-MAR-1997; SUNTORY LTD
COMMENT	OS Gentianatriflora PN JP 1997070290-A/1 PD 18-MAR-1997 PR 30-JAN-1996 JP 1996046534 PF 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI PI ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHIRO. PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAHI PC C12N15/09, A01H1/00, C07H21/04, C07K14/42, C12N9/10, (C12N9/10, PC C12R1:865), PC (C12N9/10, C12R1:19); CC strandedness: Double; CC topology: Linear; FH key Location/Qualifiers

FT	source	1. .1703	/organism='Gentianatriflora'	
FT			/variety='japonica'	
FT			/tissue_type='petal'	
FT			/clone='pgAT4'	
FT	CDS	6. .1415		
FT			/product='acyltransferase precursor'	FT
mat_peptide		9. .1412	/product='acyltransferase'	
FT				
FEATURES	Location/Qualifiers			
source	1. .1703		/organism='unidentified'	
	/db_xref='taxon:32644'	482 t		
BASE COUNT	512 a	350 c	359 g	482 t
ORIGIN				
Query Match	100.0%;	Score 1703;	DB 10;	Length 1703;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1703;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	TCATTATGGAGCAATCCAAATGCTGAAGTTCCTTGAATAATGCCAAGTTACACCCACAT	60	
Db	1	TCATTATGGAGCAATCCAAATGCTGAAGTTCCTTGAATAATGCCAAGTTACACCCACAT	60	
Qy	61	CTGACACACAGATGTCGAGTTATCGCTACCGGTAAACATTCCTCGATATCCCTGGTTC	120	
Db	61	CTGACACACAGATGTCGAGTTATCGCTACCGGTAAACATTCCTCGATATCCCTGGTTC	120	
Qy	121	ACTTGAATAGATGTCGAGTCCCTCTGTTTACGACTTCGTCACCCCAAGAACACATTTCT	180	
Db	121	ACTTGAATAGATGTCGAGTCCCTCTGTTTACGACTTCGTCACCCCAAGAACACATTTCT	180	
Qy	181	TGGACACTGTTATCCCTAACTCTTAAGCCCTCTTGTCTCTCACTCTAAACACACACGTTTC	240	
Db	181	TGGACACTGTTATCCCTAACTCTTAAGCCCTCTTGTCTCTCACTCTAAACACACACGTTTC	240	
Qy	241	CGCTTAGCGAAATTTGTTGATCCGATCAAAATCGGCGAAATCCCGAAAGTTTCAGTACT	300	
Db	241	CGCTTAGCGAAATTTGTTGATCCGATCAAAATCGGCGAAATCCCGAAAGTTTCAGTACT	300	
Qy	301	CCCGTGATGAGGGGACATCGATACTTTGATCGTTGGGAGTCTGACCAAGATTTTGACT	360	
Db	301	CCCGTGATGAGGGGACATCGATACTTTGATCGTTGGGAGTCTGACCAAGATTTTGACT	360	
Qy	361	ACCTTAAAGTTCATCAACTGCTAGATTCCTCAATGATTTGCATGGCTTTTATGTTATGC	420	
Db	361	ACCTTAAAGTTCATCAACTGCTAGATTCCTCAATGATTTGCATGGCTTTTATGTTATGC	420	
Qy	421	CACGGGTTAAGGACCATGCAAGACTATAAAGTGATCCCGCTCGTAGCCGTGCAAGTAA	480	
Db	421	CACGGGTTAAGGACCATGCAAGACTATAAAGTGATCCCGCTCGTAGCCGTGCAAGTAA	480	
Qy	481	CCGTTTTCTTACCGTGGCATACCGTGGCTCTGAGCGGCACATCATTCAGTCAAGATG	540	
Db	481	CCGTTTTCTTACCGTGGCATACCGTGGCTCTGAGCGGCACATCATTCAGTCAAGATG	540	
Qy	541	CTAAAGTTTGTAAATGTTCAATCAATGCTTGGGCTATATTAACAAATTTGGGAAGACG	600	
Db	541	CTAAAGTTTGTAAATGTTCAATCAATGCTTGGGCTATATTAACAAATTTGGGAAGACG	600	
Qy	601	CGGACTTGTGTCGCGCAATCTTCTTCCATCTTTTCGATAGATCATTAATCAAGATCTGT	660	
Db	601	CGGACTTGTGTCGCGCAATCTTCTTCCATCTTTTCGATAGATCATTAATCAAGATCTGT	660	
Qy	661	ATGCCCTAGAGGAACAATTTTGGAAACGAATGCAAGATGTTCTTGAATGTTCTCTAGAT	720	
Db	661	ATGCCCTAGAGGAACAATTTTGGAAACGAATGCAAGATGTTCTTGAATGTTCTCTAGAT	720	
Qy	721	TTGGAAGCAAAACCCCTCGATTCACCAAGGTACGAGCTACATATGTCTCTCCCTTCTG	780	
Db	721	TTGGAAGCAAAACCCCTCGATTCACCAAGGTACGAGCTACATATGTCTCTCCCTTCTG	780	

Db 58 CGTTATCTCTTACTTTCTTTCGACATTACTTGGCTACTCTTCCCTCGGTCACCATCTTT 117

QY 145 TGTCTTTACGACTTTCGTCACCAAGAACACATTTCTTGGACACTGTATCCCTTAATCTTA 204

Db 118 TCTTCTATGACTTTCACATCTTAATCCATTTTCATGGACACTATTTGTCGCCAGGCTAA 177

QY 205 AGCCCTTTTGTCTCTCACTCTAAAACTACATCTCCGCTTACGGGAAATTTCTTGTATGC 264

Db 178 AACAACTTTATCGGTCACCTCTTCAACATTTTTCCTCGTTGCTAGTAATTTGATGTAT 237

QY 265 CGATCAA-----ATCGGGGAAATGCCGAAGTTTTCAGTACTCCCGTGTAGAGGGG 315

Db 238 TTCTTAACACATGATGTTGGGTTTAAATAAAAAACAGAAATAAACACAGCTTGAAGGTG 297

QY 316 ACTCGATAACTTTGATCGTGTGCGGAGCTGACACGAGGATTTGACTACCTTAAAGGTCA 375

Db 298 ATTCTGTGTGGTTACTTTTGCAGAAATGTTGCTTCACTTAATAATTTGACAGGAATC 357

QY 376 AACTGGTAGATTCGAATGATTTGCATGGCTTTTATATGTTATGCCACGGTTTATAAGGA 435

Db 358 ATCCTCGAAAATGTGAACACTTTTATCCACTTGTACCTTCATTG---GGAATGCAATCA 414

QY 436 CCATGCAAGACTATAAGTATGATCCCGCTCGTAGCCGTGCAAGTAAACGGTTTTCCTTAAC 495

Db 415 AATTATGTGATGCGTCACGGTCCCACTTTTTCACTTCAAGTGACGTTTTCCTCGGGCT 474

QY 496 GTGGCATAGCGGTGCTGACGGCACATTCATTCAATTCGAGATGCTAAAAGTTTGTAA 555

Db 475 CGGTATATCACTAGGAATGACGAATCATCATAGCCCTTGGTGACGTAGCACCGGTTCA 534

QY 556 TGTTCATCAATGCTTGGGCTATATTAAACAAATTTGGGAAAGACGGGACTTCTTGTCCG 615

Db 535 ACTTTTGAAGGGTGACCTTCGATTATTCATCTGCTGTAGATCGTCTTTTAAACGA 594

QY 616 CGAATCTTCTTCATCTTCGATAGATGATCAATCAAGATCTGTATGGCTAGAGAA 675

Db 595 AAGGATCTCCACCGTTTGTATGATGATTGATTAACA-----TCCACATTTTAGATGA 646

QY 676 CATTTTGAACGAAATGCAAGATGTTCTTGAATGTTCTAGATTTGGAACCAACCCC 735

Db 647 AATAAGTTGAGACATACAG-----GCTCGAAAGTTTATATAACCTTCGAGCCTTGTG 702

QY 736 CTCGATTCAACAGGTACGAGCTACATATGCTCTCTCTCTTGTGAAATCCGAAAGCTAA 795

Db 703 GTCCCACTGATAAGTTGCTCAACGTTTGTGTGACCCGCAACTAATATCAATCTACTAA 762

QY 796 AGAACAAAGTACTGAATCTCAGAGGATCGGAACCGACAATACGTGTAACGACGTTCCACA 855

Db 763 AGAAAAGGTCTTAA-----CCCAAGTCCCAAACTTGGAGTACATGT--CATCTTTTACGG 816

QY 856 TGACGTGTGGATACGATGACATGTCATGTCATAATCAAAAGATGACGTCGTATCAGAG 915

Db 817 TAACCTTGTGTTATATGAGTTGATAGCTAGCGAATCA-----CTCGTAAAAATAG 867

QY 916 AATCATCGAACGACGAAAATGAGCTCGAGTACTTTCAGTTTACAGCGGATTCGCCAGGAC 975

Db 868 GAGAAAGAAAGGCGAAGACGAGTTAGAACAGTTTCAATATCAACCATTTGATGTCGATCTC 927

QY 976 TTTCTAGCGCCCGTGTCCGCTTAATCTTTGGCAACTGTCTTGGCTGATCGGTTGCA 1035

Db 928 GTCTGATCCACCAATTCGCACAGCCTACTTTTGGTAACTGTGGTGACCATGTTGTC 987

QY 1036 AAGCAACATAAAGATTTAGTTGGGGATAAAGGGCTTCTTGTGAGTTGCAGCTATTG 1095

Db 988 CTTAAAAAATGCTGTTTGTACTAGCGAAATGGGTATGCACTTGTGCTAAAGTAATTG 1047

QY 1096 GAGAGCCATTCGAAAGAGGTTGCACAACGAAAAGCGGTTCTTGCAGATGCAAAAACCTT 1155

Db 1048 GAGAGCTATATGCAAAATGATATATAATAAGGACCGAATCTTGAAGATGCCCGAGAT 1107

QY 1156 GGTATCGGAATCTAATGGAATCCCTTTCAAAAAGATTTCTCGGGATTAACGGATCGCC 1215

Db 1108 GG---CATGAACCTTTTCATGATCCCGGCTAGGAAGATT---GGTGTGCTGGTACACCTA 1161

QY 1216 AGTTTCGATTGCTATGTTAGATTTTGGATGGGAAAGCCTGCAAAATTTTGACATTACCT 1275

Db 1162 AGCTCAACTGTCAGACTTTTGATTTGGTGGGGAAGCCGATAAAGTATGAGACTGTTT 1321

QY 1276 CTGTTGATTATGCAGAAATTTGATTTATGTTTCAGTCCAGGGATTTTGA AAAAGGTGTG 1335

Db 1222 CAATAGACTATAATACGTCGATTTCTATTAATGCAAGCAAAACATCAGCACAAAGATCTG 1381

QY 1336 AGATTGGAGTATCATTTGCCCTAAGATTCAATGATGATGATTTGCAAAAATCTTTGAAGAAG 1395

Db 1282 AATTGGATTGAGTCTACCGAGTATGCAATGAGGCGGTTTCTAGCATCTTTTGATGAAG 1341

QY 1396 GCTT 1399

Db 1342 GATT 1345

RESULT 4

E12754

LOCUS E12754 1622 bp DNA PAT 24-JUN-1998

DEFINITION Gentianatriflora mRNA for acyltransferase, complete cds.

ACCESSION E12754

VERSION E12754.1 GI:3251586

KEYWORDS JP 1997070290-A/2.

SOURCE unidentified.

ORGANISM unclassified.

REFERENCE 1 (bases 1 to 1622)

AUTHORS Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y., Yonekura,K., Mizutani,M. and Kusumi,T.

TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY

JOURNAL Patent: JP 1997070290-A 2 18-MAR-1997;

COMMENT SUNTORY LTD

OS Gentianatriflora

PN JP 1997070290-A/2

PD 18-MAR-1997

PF 30-JAN-1996 JP 1996046534

PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI

ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHIRO,

PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PC

C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC

C12R1:R65).

PC (C12N9/10,C12R1:19);

CC strandedness: Double;

CC topology: Linear;

FT Key Location/Qualifiers

FT source 1..1622

FT /variety='japonica' /organism='Gentiana triflora' FT

FT /tissue_type='petal' /clone='pgAT106'

FT CDS 35..1474

FT /product='acyltransferase'.

FEATURES

source

1..1622

/organism='unidentified'

/db_xref='taxon:3264'

BASE COUNT 478 a 345 c 322 g 477 t

ORIGIN

Query Match 11.1%; Score 189.2; DB 10; Length 1622;

Best Local Similarity 52.2%; Pred. No. 1e-37;

Matches 728; Conservative 0; Mismatches 618; Indels 48; Gaps 12;

QY 25 TGAAGGTTCTTGA AAAATGCCAAGTTTACACCACCATCTGCACACACAGATGTCAGTTAT 84

Db 57 TCAAGGTTCTTGA AAAATGCCGTTTTCGCCACCACCGGAC---GCCGTCGCCGAGTTTA 113

QY 85 CGCTACCGGTTAACATTTCTTCGATATCCCTCGTTGTCACCTTGAATAAGATGCAGTCCCTTC 144

Db	1179	ATTGGATATCGCCGCCCGATCAACATCTGCGGCCAACAGTTCGACGCTCATTTATGTCG	1238
Qy	1198	GGATTACCGGATCGCCCTAAAGTTCCGATTTCGATTGCTAGATTTTGGATGGGGAAGCGTG	1257
Db	1239	TTGGATCGGCACACGCAATCTGCATGATTTTGATGCAGATTTTGGTTGGGAAGCGTTG	1298
Qy	1258	CAAAATTTGACA-----TTACCTCTCTGTTGATATGACAGAAATGATTTGATTCAGTCCA	1314
Db	1299	AAAAGCATGAATCTGTTTCAACTAATCCTTCGGCAACACATAAATTTTTCATCTCTCGGTCCA	1358
Qy	1315	GGGATTTTGAAGAGGTGTGGAGATTTGGAGATATCATTTGCCTAAGATTCATATGATGCAT	1374
Db	1359	GAGATTTAAAGGAGCACTTGAGCTTGGCATTTCTTTGCCCTAAGAATAGGATGACGCAT	1418
Qy	1375	TTGCAAAAATCTTT	1388
Db	1419	TTGCCACCACTTTT	1432
RESULT	5		
LOCUS	AB026494	1622 bp	PLN
DEFINITION	Gentiana triflora GAT106 mRNA for acyltransferase homolog, complete cds.		
ACCESSION	AB026494		
VERSION	AB026494.1	GI:7415596	
KEYWORDS	acyltransferase homolog.		
SOURCE	Gentiana triflora CDNA to mRNA.		
ORGANISM	Gentiana triflora		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.		
AUTHORS	Yonekura-Sakakibara,K., Tanaka,Y., Fukuchi-Mizutani,M., Fujiwara,H., Fukui,Y., Toshihiko,A., Tamauchi,M. and Kusumi,T.		
TITLE	Molecular cloning and biochemical characterization of hydroxycinnamoyl-CoA:anthocyanin 3-O-glucoside-6-O-hydroxycinnamoyltransferase from Perilla frutescens and diverse plant acyltransferase homologs Unpublished (1999)		
JOURNAL	2 (bases 1 to 1622)		
AUTHORS	Tanaka,Y. and Yonekura-Sakakibara,K.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-APR-1999), to the DDBJ/EMBL/GenBank databases. Yoshikazu Tanaka, Suntory Ltd., Institute for Fundamental Research; Wakayama-dai 1-1-1, Shimamoto, Osaka 618-8503, Japan (E-mail:Yoshikazu.Tanaka@suntory.co.jp, Tel:81-75-962-8807, Fax:81-75-962-8262)		
FEATURES	Location/Qualifiers		
source	1..1622		
gene	/organism="Gentiana triflora"		
CDS	/db_xref="taxon:55190"		
	35..1474		
	/gene="GAT106"		
	35..1474		
	/gene="GAT106"		
	/codon_start=1		
	/product="acyltransferase homolog"		
	/protein_id="BAA93452.1"		
	/db_xref="GI:7415597"		
	/translation="MAGNSDIDKVKRCRVAPPDVAETVPSLFFDMRWLISDAEHLHFYRFRPCPSKFTIAGHSKSLVLKHLPLAGNLIPVDSDDRMPELRYKKGS SVLSLTAEQSMDFLAFODHORYKFNDLIPOLPEIYVTSGEDVLPALQVTVPSN TGCIGRNHLQVLGDASSFLHPKMLVLDKNSGDSLKFLPSSLPMYDSVVODPFH IRRKIYNERLLKSQGTPTVLPNPAISKDEVRAATFILHPIDIMKKKFIISKNNRLTGS SNNYITFTSALINTWCLSDTIVREKVEDAAHNLCAFINCRFAPPIPQNY FGCNIVPCVMGSTEHOELVNGEGLSVAATGDAIHKRLHDYEGILRGDWISPPRSTA APSTLIYVVGSAQRNVHDFDADFQNGKLEKHESVSTNFSATLILISRRRFRKGRSL		
BASE COUNT	478 a	345 c	322 g 477 t
ORIGIN			

Query Match		11.1%	Score 189.2;	DB 12;	Length 1622;
Best Local Similarity		52.2%;	Pred. No. 1e-37;		
Matches 728;		Conservative 0;	Mismatches 618;	Indels 48;	Gaps 12;
QY	25	TGAAGGTTCTTGAANAATGCCAAGTTACACCACCACCTCTGACACAACAGATGTCGAGTTAT	84		
DB	57	TCAAGGTTCTTGAGANAATCCGTTGTCGCCACCACCGGAC---GCCGCGCGGAGTTTA	113		
QY	85	CGCTACCGGTAACTCTTCGATATCCCTGGTTGCACTTGAATAAGATGCAAGTCCCTTC	144		
DB	114	CAGTCCCACTGTGTTTTTCGACATCGGATGGTTGATCTCTGATGACAAACACCATCTGC	173		
QY	145	TGTTTTACAGCTTTCGGTACCCCAAGAACACATTTCTTGGACACTGTTATCCCTCAATCTTA	204		
DB	174	ATTCTACAGATTCCGCCATCTCTGTCACAACT---CTCTAAATTTATCATTTTCATCCATTA	230		
QY	205	AGGCCTCTTGTCTCTCACTCTAAACACACTAGTTCGGCTTAGCGGAATTTGTTGATGC	264		
DB	231	AATCGTCCCTTTCCCTTGTCTCAACACACTTCTTCGGTAGCCGGAATTTGATTTGGC	290		
QY	265	CGATCAAAATCGGGCGA---AATGCCCAAGTTTCAGTACTCCCGTGATAGGGCGACTCGA	321		
DB	291	CGGTAGATTCTCCGATAGATGCCGGAGTTGGTTAC-----AAGAAGGGGACTCCG	344		
QY	322	TAACTTTGATCGTTGGGAGTCTGACAGGATTTTGACTACCTTAAAGGTCAATCAACTGG	381		
DB	345	TTTCTTTTAACAATTGCAGATCGAGCATGGATTTTGATATCTCGCGCGAGATCATCAGA	404		
QY	382	TAGATTCCAATGATTTGCAATGGCCTTTTTTATGTTATGTCACGGGTTTATAAGGACCATGC	441		
DB	405	GGGATTTCTTATAAATTAACAGATTTGATTTCCGCACTGTCAGAACCGATTTGTAACC---T	461		
QY	442	AGACTATAAAGTATCCCGCTCGTAGCGTGCAAGTAAACCGTTTTCCTCAACCGTGGCA	501		
DB	462	CGGCGCAAGATATTACCACTTTTTCGCTTTACAGTGACGGTGTCTTCCCAACCCGGTA	521		
QY	502	TAGCCGTGGCTCTGACGGCACATCATTAATTCAGATGCTAAAGGTTTTTGTAAATGTCA	561		
DB	522	TATGCAATTTGGACGCAATCTTCATCAAGTTCTTGGTGATGCCAGTTCTTTCTGCAATTTTA	581		
QY	562	TCNAATGCTTTGGGCTATATTACAAAT---TTGGGAAAGACGGGACTTGTGTGTCGCGCA	618		
DB	582	ATAAATATTATGGGTTTGGTTGACAAATCCAAATCCAAATGGAGATTCATTAAGTTCCCTTCCACTT	641		
QY	619	ATCTTCTTCCATCTTTCAGATAGATCAATCAAGATCTGTATGCTCTAGAGGAAACAT	678		
DB	642	CTTCTCTACTATGTACGACAGATCTGTGGTGCAGATCCATTTTCATATTCGTCGAAATA	701		
QY	679	TTTGGAAAGAAATGCAAGATGT---TCTTGAATGTTCTCTAGATTTTGGAAAGCAAAACCCC	735		
DB	702	TCTACAATGAAGAAAATGCTCAAAATCTCAGGCGCACACTACTGTTCTATAATCCAGCAA	761		
QY	736	CTCGATTCAACAAGGTACGAGCTACATATGCTCTCTCCCTTGGTGAATCCAGAAGCTAA	795		
DB	762	TTTCTAAAGATGAAGTTCGAGCACCTTTCATCTACACCTTATGATCATCATGAAGCTCA	821		
QY	796	AG-AACAAGTACTGAATCTCAGAGATCCGAACCCACAAATA-----CGGTAAACGA	846		
DB	822	AGAAATTCATTTCTGCTCAAAAATTCGCAACTTAACCGGTAGTAGTAATTAATCTGTCAA	881		
QY	847	CGTTCAATAGCGTGTGATAGTATGGACATGATGCTCAATATGCTCAATCAAAAGATGAGTGC	906		
DB	882	CTTTCACGGTGACATCTGCACATGATCTGGACATGCTTGTGCAATCATTTAGACACCGTCG	941		
QY	907	TATCAGAGGATCATCAGACGCAAAATGAGCTCGAGTACTTCAGTTTACAGCGGAT	966		
DB	942	TAAGAGAGAA---GGTGAAGAGGATTAACATGACACAACTTATGTGCTTTCATCAACT	998		
QY	967	GCGGAGGACTTCTGACGCCCGCTGTCGCGCTAACTACTTTTGCAACTGTCTTTCGCTCAT	1026		
DB	999	GCGGACAGCTTTTGTCTCCGCCCATACCTCAAAATTTACTTTTGGAAATTCATAGTGCCTT	1058		
QY	1027	GCCTTGCAAAAGCAACACATAAAGATTAGTTGGGGATAAAGGCTTCTTGTGTCAGTTG	1086		

Db	1059	GTATGGTGGGATCGACTCATGAGCAACTTGTAGAAATGAAGGTTGTCTGCTAGCTGCAA	1118
Qy	1087	CAGCTATTGGAGAGCCATTGAAAGAGGTTGCACACGAAAGAGGGTTCTTCGCAGATG	1146
Db	1119	CCGCCATCGGAGATGCTATCCATAAGAGGTTTACATGACTACGAGGAAGTAATCTCGAGAGGAG	1178
Qy	1147	CAAAAATCTGGTTATCGGAATCTA-----ATGGAATCCCTTCAAAAAGATTTCTCTCG	1197
Db	1179	ATTGGATATATCGCCGCCCGCATCAACATCTCGGGCACCAGGTCGACGCTCATTTATGTCTG	1238
Qy	1198	GGATTACCGGATCGCCTAAGTTCGATTCTGATGCTAGATTTTGGATGGGCAAGCGCTG	1257
Db	1239	TTGGATCCGCACACGCAATGTCATGATTTTGTATGACGATTTTGGTTGGGCAAGCTTG	1298
Qy	1258	CAAAATTTGACA---TTACCTCTGTTGATTATGACGAATTTGATTTATGTATTCAGTCCA	1314
Db	1299	AAAGCATGAATCTGTTTCAACTAATCCTTCGGCAACACATAATTTTGATCTCTCGGTCCA	1358
Qy	1315	GGGATTTTGAAGAGGTGTGGAGATTGGAGTATCATTTGCCCTAAGATTCATATGGATGCAT	1374
Db	1359	GAAGATTTAAAGGAGCACTTGAGCTTGGCATTTCTTTGCCCTAAGAATAGGATGACGCAT	1418
Qy	1375	TTGCAAAAATCTTT	1388
Db	1419	TTGCCACCATTTT	1432
RESULT 6			
E12756			
LOCUS			
E12756		1479 bp	DNA PAT 24-JUN-1998
DEFINITION			
Perilla ocimoides mRNA for acyltransferase, partial cds.			
ACCESSION			
E12756			
VERSION			
E12756.1 GI:3251568			
KEYWORDS			
JP 1997070290-A/4.			
SOURCE			
unidentified.			
ORGANISM			
unclassified.			
REFERENCE			
1 (bases 1 to 1479)			
AUTHORS			
Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y., Yonekura,K., Mizutani,M. and Kusumi,T..			
TITLE			
GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY			
JOURNAL			
Patent: JP 1997070390-A 4 18-MAR-1997; SUNTORY LTD			
COMMENT			
OS Perilla ocimoides			
PN JP 1997070290-A/4			
PD 18-MAR-1997			
PF 30-JAN-1996 JP 1996046534			
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI			
ASHIKARI TOSHIHIKO. TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHIRO,			
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PI			
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC			
C12R1:865),			
PC (C12N9/10,C12R1:19);			
CC strandedness: Double;			
CC topology: Linear;			
FH Key			
Location/Qualifiers			
FT source			
1..1479			
/organism='Perilla ocimoides' FT			
/tissue_type='leaves'			
FT CDS			
1..1343			
/product='acyltransferase'.			
FT			
Location/Qualifiers			
1..1479			
/organism='unidentified'			
/db_xref='taxon:32644'			
BASE COUNT			
420 a 316 c 331 g 411 t 1 others			
ORIGIN			

Query Match		10.1%; Score 172.2; DB 10; Length 1479;
Best Local Similarity		49.5%; Pred. No. 2.2e-33;
Matches		701; Conservative 0; Mismatches 668; Indels 48; Gaps 8;
Qy	78	GAGTATACGCTACCGGTAAACATCTTCGATATCCCTGGTGGTACCTTGAATAGATGACG 137
Db	48	GAGCAATCGGTGCGCTCACATCTTCGACATGAGCTGGCTGCATTTTCATCCCATGCTT 107
Qy	138	TCCCTTCTGTTTACGACTTCCGTAACCAAGACACATTTCTTGGACACTGTTATCCCT 197
Db	108	CAGCTCCTCTTTCAGCAATCCCTTGTTCAGCAACATTTTTCAGATCCATCGTTCCA 167
Qy	198	AATCTTAAGGCTCTTCTCTCACTCTAAACACACTACGTTCCGCTTAGCGAAATTTG 257
Db	168	AACTCAACAATCTCTCTTAAACATCTACACTTCTTCCCTCTCATGCAATTTA 227
Qy	258	TTGATGCGGATCAATCGGGGAAATCGCGAAGTTTTCAGTACTCCCGTGATGAGGGGAC 317
Db	228	ATCTACCCCTTCATCCCGGAGAAATGCGGAGTTTCGGTA-----TCTATCCGGGAC 281
Qy	318	TCGATAACTTTTGATCGTTCGCGAGCTGACCCAGGATTTTGACTACCTTAAAGTCAAC 377
Db	282	TCGGTTCTTTTACCATCGCAATCTTAGCCAGGACATTCGATGATCTCGTCCGAATCGT 341
Qy	378	CTGTAGATTTCAATGATTTGCATGGCCCTTTTATGTTATGTCACCGGTTTATAGGACC 437
Db	342	CCAGATCTCCCGTTAGGCTCTACAACCTTGTCCCTAAATTCGCCCATTTGTCGAAG 401
Qy	438	ATGCAAGACTATAAGTGATCCCGCTGATCGGTGCAAGTAACGGTTTTCCTTAACCGT 497
Db	402	---TCCGATAGAAACTCTTCCAGTTTTCGCGGTGCGAGTGACTCTTTTCCAGGCGGA 458
Qy	498	GGCATAGCGTGGCTCTGACGCGACATCATCAATTCGAGATGCTAAGATTTTCTAATG 557
Db	459	GGCGTGTATTGGAATAGCAACGATCACACCGTTACGACGCGCCCTGCTTCTCGCG 518
Qy	558	TTTCATCAATGCTTGGGCTATATTAACAAA---TTTGGAAAGACGCGACTTGTGTC 614
Db	519	TTTATAAGGCTTGTCTTCAATGAGCAACACATTTGAAATGAAGATGAAGATGAAG 578
Qy	615	CGAATCTTCTTCATCTTCGATAGATCGATATCAAGATCTGTATGGCTAGAGAA 674
Db	579	TTTAAATCTTTCGCGAGTTTTCGATAGATCCGTCATAAATATCGCAGCAAAATTTGACTCC 638
Qy	675	ACATTTTGAACGAATGCAAGATGTTCTTGAATGTTCTAGATTTGGAGCAAAACC 734
Db	639	ATTATTGGAGAAACGCGCTAAAATTTTCT-----TTGCAATCTCGTCATCCCTCA 689
Qy	735	CCTCGATTCAACAGGTACGAGCTACATATGCTCTCCCTTGTGTAATCCAGAAGCTA 794
Db	690	TTACCGAGGACCGCATTCGAACACGTTCTGTTTTCACCCCAATCCAAATTAAGAAATG 749
Qy	795	AAGAACAAAGTACTGAATCTCAGAGGATTCGAACCGAATACGTGTAAACGAGTTTCA 854
Db	750	AAG-----GGTTGATTCAGTCCAGAGTTCCAAAGTTTAGTCCATCTCTCATCTTTGTA 803
Qy	855	ATGAGCTGTGATACGATGACATGATGCTAATCAAAAGATGAGCTCGTATCAGAG 914
Db	804	CGGATTCGAGCTTATATGTGGCTGGCGATACAGAAATCATTCACAGCAGATGAAG----- 858
Qy	915	GAATCATCGAACGAGAAATGAGCTCGAGTACTTCAAGTTTACAGGGGATTCGCGAGGA 974
Db	859	-----ACCAAGACACGAGATGATCATTTTCTTGATTCGGTTCGATCTAAGGCCA 908
Qy	975	CTTCTGAGCCCCCGTGTCCGCTTAATCTTTCGCAATGCTGTTCGGTCTAAGCTTGCA 1034
Db	909	CGATTAGATCGCGGGTTCCTGAAATTAATCTTTCGGAATCTTATCGTACGCGCTGCG 968
Qy	1035	AAACAACACATTAAGATTTAGTTGGGATTAAGGGCTTCTTGTTCGATGTCGACTATT 1094
Db	969	AGATTCGCGGCGGACGAGCTGCTGGGAGAGAAAGGGGTGTTTTCGGCAGCTGAGGTAA 1028
Qy	1095	GGAGAAGCCATTGAAAGAGGTTGCAACAAAGAAAGGCGTTCTTTCGAGATGCAAAACT 1154

Db	1029	CGCGCGGAGATAAAAAAAGAT---CAACGACAAGAGAATATTAGAAACGGTGGAGAA 1085
Qy	1155	TGGTTATCGGAATCTAATGGAATCCCTTCAAAAAGATTTTCGGGATTTACCGGATCGCT 1214
Db	1086	TGGTCGCGGAGATTCGTAAGCGTTGCAGAAATCATATTTTTCGGTGCAGATCGAGC 1145
Qy	1215	AAGTTCGATTCGATGTTAGATTTGGATGGGAAGCCCTGCAAAATTTTGACATTACC 1274
Db	1146	AAGCTAGATCTTTACGGTGCAGATTTTGGATGGGAAGCGGAGAAAGCAAGAAATATTG 1205
Qy	1275	TCTGTGATTTGACAGATTG---ATTATGTGATTCAGTCCAGGATTTTTCGAAAAGCT 1331
Db	1206	TCGATTCGAGGAGAAATATGCAATGACRCTTCTGAAGCCAGGATTTTTCGAAGGAGA 1265
Qy	1332	GTGAGATTCGAGTATCATTCCTTAAGATTCATATGGATGATTTTGCAAAATCTTTGAA 1391
Db	1266	TTGGAGTTTCTGCTTTCCTTAAGGACAAATTTGCTGCTTTTTCGCTGCTATTTTCA 1325
Qy	1392	GAAGCTTTTCTGCTCTGCTATAGTCTCTTTAATAGAACCATATTTTCTGCTGCAATAAGTA 1451
Db	1326	CTGGAATTAATGTTAATAATGATGATTAATACTAATTAATTAATTAATTAATTAATTA 1385
Qy	1452	CCAAGTCTTTTAGTAACACTACACCAACCCCTACTTT 1488
Db	1386	TTAAGTGTGAGTAACGTAAGATAATATCCCTATTAT 1422
RESULT 7		
AB029340		
LOCUS		
DEFINITION		
AB029340 1476 bp mRNA PLN 04-APR-2000		
Perilla frutescens mRNA for anthocyanin acyltransferase, partial cds.		
ACCESSION		
AB029340.1 GI:7415645		
VERSION		
Perilla frutescens leaf cDNA to mRNA.		
KEYWORDS		
Perilla frutescens		
SOURCE		
ORGANISM		
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Lamiales; Lamiales; Perilla.		
REFERENCE		
1 (sites)		
AUTHORS		
Sakakibara, K.Y., Tanaka, Y., Mizutani, M.F., Fujiwara, H., Fukui, Y., Ashikari, T., Yamaguchi, M., and Kusumi, T.		
TITLE		
Molecular and biochemical characterization of a novel hydroxycinnamoyl-CoA: anthocyanin 3-O-glucoside-6		
JOURNAL		
Unpublished (1999)		
REFERENCE		
2 (bases 1 to 1476)		
AUTHORS		
Sakakibara, K.Y. and Tanaka, Y.		
TITLE		
Direct Submission		
JOURNAL		
Submitted (24-JUN-1999) to the DDBJ/EMBL/GenBank databases. Keiko Y Sakakibara, Sutory Research Center, Fundamental Research, Plant Biotech; Shimamoto-cho, Wakayamada, 1-1-1, Mishima-gun, Osaka 618-8503, Japan (E-mail: Keiko.Sakakibara@suntory.co.jp, Tel.: +81-75-962-8807, Fax: +81-75-962-8262)		
FEATURES		
source		
1..1476		
/organism="Perilla frutescens"		
/db_xref="taxon:48386"		
/tissue_type="leaf"		
<1..1343		
/codon_start=3		
/product="anthocyanin acyltransferase"		
/protein_id="BA093475.1"		
/translation="VIEICRVGPPDVAQSVPLTFDMLTWHFPMQLQLLFVFFPC SKDFSEIVPKLQSLKTLIHFFPLSCNLIYVSPSPKPEFRLYGLSDVSFTIAGS SDFDLDVGNRPSPVRLYFVPLPPIVEESDRKLFQVAVOVLPLPGRGVGIAT HHVSDAPSLAFITANSSSKHIENEDDEEPSKSLPVFORSVIKYKTFKDSIWRNA LKPLPSRHPSLPTDRITFTTFTQSKIKKLKQVRSVLSVHLSFSVAIAAYMWAG IKSFTADEQDQDNEAFLLPDLRPLDPVPENFGNLSLALPMRRELVGEG VFLAAEIVIAIEIKRKIRILETKWSPKIRKALQKYSFVAGSKSLDLYGDFGW GKARKQELISIDGKVMTLCKARDFEGGLEVLCLSLPKDKMDAFAAYFSLING"		
CDS		

Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremetska,I., Kim,C., Lenz,C., Li,J., Liu,S., Luoro,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
 Direct Submission
 Submitted (30-JAN-1999) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA

TITLE JOURNAL

COMMENT

On Dec 30, 1998 this sequence version replaced gi:2734094.
 Bases 1-9262 of clone F21M11 overlap with bases 68998-78259 of
 'TAMU' BAC clone F20D22 (AC002411) and bases 119525-119914 of clone
 F21M11 overlap with bases 1-389 of 'TAMU' BAC clone F21B7
 (AC002560).

e-mail for correspondence: arabidopsis.stanford.edu
 Genes with similarity to proteins in the databases are described as
 'putative', '-like' or 'similar to'. Genes that have EST
 similarity but no significant protein similarity are described as
 'unknown proteins'. Genes that are annotated based only on gene
 prediction software are described as 'hypothetical proteins'.
 The software programs used to predict genes include: Grail
 (informatics.group, Oak Ridge National Laboratory,
<http://compbio.ornl.gov/section/index.html>), GENSCAN (Chris Burge,
<http://gnomic.stanford.edu/~chris/GENSCANW.html>), Fexa (V.Solovyev
 & A.Salamov, Sanger Centre, <http://genomic.sanger.ac.uk/>), and
 NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, <http://www.cbs.dtu.dk/NetPlantGene.html>).

FEATURES

source

1..119914
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="I"
 /clone="F21M11"
 1..7000

misc_feature

/note="Proteins in this region are annotated in the F21M11
 entry, AC002411."
 1..9262

misc_feature

/note="overlap with bases 68998-78259 of 'IGF' BAC clone
 F21M11."
 7684..10913

gene

/gene="F21M11.1"

CDS

join(7684..7866,7965..8027,8122..8160,8250..8381,
 8458..8728,8875..9893,10003..10256,10344..10410,
 10501..10551,10685..10781,10885..10913)
 /gene="F21M11.1"

/note="Hypothetical protein"

/codon_start=1

/protein_id="A010665.1"

/db_xref="GI:4204284"

/translation="MAPNLRIRKACDAMKLLGISETKTKRAFLRKLKTYENNWDFTIEE
 DAYKVLDAIFDADAQSTKKNKKEKKKKEKKSRSVATSRGRKKAPELQVDEE
 DMDEDFPLKRLRRGRASSSSSSSYNNEDLKTOPEEDEDGVTPLPLAKRY
 VRNGEGLATVYNNASPSRSLRSMEEVPPVLLPAHPMETKVSASALVILND
 EPNIDHKPVSDGNCASPMLEKGNHIVQEMDWETKDLINDTAMVSPSAIGES
 SEHKVAASVELASSTSGEAKICLSPATGETTNLHPSMEDLRAMEEKLKYSKI
 VHPFSLVGMKSCYCIDLAKNSTQLLETETVCDMSKAGDEAGVIGSMPLVVP
 ECISGDGKWAISNMKDIITAGEENVEIPWNEINERKVPFRYPHPSFVODAPVIFS
 LSSFDSQSCSTSCIEDCLASEMNCNCAIGVDNGEAVTLDDLKLEFLEAREADQ
 RKQVLRFCBCEPLERAKKVEILEPCGHLKRGAIKCEWPKCGCTKRCGRNVYQRMHN
 KLOVFTPNKGWGLRLEKLPKGAFCICEYIGELITPELYQRSFEDKPLPVLDAH
 WSGEERLEGDLCLDGMFYGNISRLNHRCLDANLITPVOVETPDQHYLAFFT
 RDTAEAMEALWDYGDIDFNNDMSLMKPFCDCLGSRFCRNKRKSTKQIILNKA"
 join(13033..13488,13717..13926,14139..14514)
 /gene="F21M11.2"
 13033..14514
 /gene="F21M11.2"

mRNA

join(13061..13488,13717..13926,14139..14316)
 /gene="F21M11.2"

CDS

/note="Similar to acid phosphatase; Location of ESTs
 110C2T7, gb|T42036, and 110C2XP, gb|A1100245"
 110C2T7
 /codon_start=1
 /protein_id="A010666.1"
 /db_xref="GI:4204285"

mRNA

gene

CDS

/translation="MDRTWFLSLTIAASLLVGVVSGADMNLLNQLRGLGSSSSONGIVS
 KGITLCKGYCESRINVEHNRIKDVVQECVSHKIDMTSSQKDDVARTVDEK
 LHFSGMCCSKGCDGMADTIFDIDTLLSTIPYHKKNGFGGKSLNSTFQVIVQYI
 APAVPMKLYLHDIRERGIKFLISSRKVELRSATVDNLIOAGYVGNLSLIVYA"
 OREVKYKSEKKRMLSLGVRVWGMGDSSWSPAGCPLPRRTFKLPSNLIYYA"
 complement(join(14004..14312,14565..14641,14787..15831,
 15906..16006,16327..16362))
 /gene="F21M11.3"

complement(14004..16362)

/gene="F21M11.3"

complement(join(14567..14641,14787..15831,15906..16006,
 16327..16362))

/gene="F21M11.3"

/note="Unknown protein; Location of ESTs 40C3TT,
 gb|AA728590 and40C3T7, gb|T04573"

/codon_start=1

/protein_id="A010667.1"

/db_xref="GI:4204286"

/translation="MLDQISGNHTHEKLSVETAEPHHLNDRVHIVEIPKASVIPITE
 ICDAEEKSPSTISKRVTFSKVKTYEHVSESEVSEELSEKNEVESEKRSLSKSK
 TDDOIIEVASNSGSGYPNHRKYNCRSDDDIEDEDFDCSDSLDEDEEYSDVQFSE
 DSLHNPTEKVYTQDIGDKTEEDSKLRRSNETVRDGNHYDQGVLPVENLITQMSAK
 SKGRTKQKQSKENSFNIAQOEKRSDFDQDDITLSVKPKCRPEPKLRQKE
 LAVDASLWLSSTESGECNSAMTTLPEKLSKSCYSKPLRINHDDRPVLCALTL
 EDIKQFASLSTPRKSPSKSPDETPIITGTWGYWGNRSKAIDCGSASFSGIPNTSSKY
 REDKSVNHSHTPEARLEKALNNIDK"
 complement(16835..20238)
 /gene="F21M11.4"

gene

CDS

complement(join(16835..17185,17274..17392,17491..17656,
 17795..17885,17982..18079,18175..18361,18504..18604,
 18705..19049,19134..19349,19439..19791,19862..19967,
 20188..20238))
 /gene="F21M11.4"

/note="Hypothetical protein"

/codon_start=1

/protein_id="A010669.1"

/db_xref="GI:4204288"

/translation="MGSSPEARAAQVPSMILFLEIITVHVYNNRRKLNRDLVLSA
 NLNIPKRVTRDLRREMSVISYKSLNAAVSVHLPOLOIPNDCNYKNALNNSPKH
 GESEDSMTDKOVSKRGSGTDSRRDGLPTSESDPRPKHODWTEKOLSDHLLVE
 FSEYDAANTPESYTEQAARVDRDITASEQPSNAARKICGDSFQESSPNPKTQDP
 TLLRLMESLRSDPTVRAQNHQWYFVFWIYISFVLCRLSRNNKRCVIFVCQQLPK
 SHTEQDSKKRRDITASDAMENHLKVPKRENNLMQKSDIDCNGKSANSDDQSEKIS
 KALEQTSNITTCGFCQCSARVSATGEMLHYSGRVPDGDIDIFRSNVIHVSACIEWA
 POVYEGEDTVKNLKAELAGMKIKCTKSLKGAALCGFKVSKRRSYHVCAREISCR
 WYDEFLLCPSHSSVYKFNKSGHRVRAEPLKINPAELCSLEOTPAFTKELVLCG
 SALSXDKKLMEASVRENATISRYNPSVTHVIASTDDEKGACTRLLKVLMLGILNKW
 IINAAWKAASKASQPVDEEPEFIQIDGCGQCPKTAARLRAETNPKLFEGLKFTFF
 GDFYKGYKEDLQNLVKVAGGTILNTEDELGAESSNNVNDORSSIVYVNIIDPPHGCAL
 GEVYTIWQRANDAEALASQTSRLVGHWTWVLESIAGYKLHPVIG"
 20752..24647
 /gene="F21M11.5"

gene

CDS

join(20752..20994,21362..21497,21596..21740,21825..21994,
 22102..22178,22542..22758,22920..23064,23343..23442,
 23599..23693,24021..24072,24227..24298,24441..24647)
 /gene="F21M11.5"

/note="Hypothetical protein"

/codon_start=1

/protein_id="A010668.1"

/db_xref="GI:4204287"

/translation="MGANSKVTASTVTIVAVFLLICGGRTAVDETEFHGDSKLSGI
 IIPGFASLQRAMSILDCPYTPLDFNPLDLVLDITKLLSANCWFKCMVLDYDNGTD
 HPCKSRDGLSAITELDPGYITGPLSTWKMELKWCVEFGIEANAIVAVDPKRLS
 PKLEERDLYFHKLKLTFTALKRGSPSIVFAHSMGNVRYFLEWLRLKLEIAPKHYL
 KWLDOHIIHAYFANGAPLLGSLVAKIKSTLSQVTEGLPVSEGTARLLNSFASLWMLPF
 SKNCKDNTFWHFGSGAAKOKRVTHCDEEVEQSKYSGWPTNIIHIEIPSTSARELA
 DGTFLKAIEDIYDPSKRLMLHUKKYVFFVIRNIAHRSSLAGLELHDDPVSFNPPLTPW
 ERPKINVCYIEAHKLEVEGYFAPSGRPYDNDWIITIIYEESGLSDVSVGTVDG
 NAGPIPTGEMTVYHSLSWCKNMLGPKVNITMAPQIILGKIQKQPEHGDGSDVHVELNVD
 HBHGSIIIANMTPKAPRVKITYFDESDSIPGKRTAVWELDKSGY"
 27777..28734
 /gene="F21M11.6"

mRNA

gene

27777..28734

/gene="F21M11.6"
28007. .28465
/gene="F21M11.6"
/note="Unknown protein; Location of ESTs 20312477,
gb|H76794 and 203124XP, gb|AA605510"
/codon_start=1
/db_xref="GI:4204289"
/translation="MNTYMLRPPRLVLTADKRERDAFISSVTDNPPPIAKFPSPPP
KLVPPVPNIPISKKSTAAAEPIGSLAGLYLSHEYLTQSTLFGQBNQARAQESS
KIKPSHTVEPEECPPKRRYREVANLURSDGQPLGVNPAQLARFLKL"
complement(29264. .32033)
/gene="F21M11.7"
complement(join(29264. .31015,31312. .31414,31484. .31587,
31782. .32033))

Query Match 6.7%; Score 114; DB 12; Length 119914;
Best Local Similarity 46.2%; Pred. No. 2.8e-18;
Matches 612; Conservative 0; Mismatches 685; Indels 27; Gaps 6;

QY 85 CGCTACCGGTAACATCTCTCGATATCCCTGGTGGCACTTGAATAAGATGCAGTCCCTTC 144
DB 46008 CTCCTGCTCACCTCTCTCGATGCCCCCTGGCTCTCTCTCCACATCGCGGATCTCTCT 46067

QY 145 TGTCTTACGACTTTCGGTACCCCAAGAACACATTTCTTTGGACACATGTTATCCCTTAATCTTA 204
DB 46068 TCTTCTTCTTACCAAACTCAACTGAATCTTTCTCTCCCAAGACTTTGTACCAACCTCA 46127

QY 205 AGGCTCTTTGCTCTACATCTAAACACTAGTTCCGGCTTACGGGAATTTGTTGATGC 264
DB 46128 AACATCTCTCTCAATCTCTCAACATTTCTTCCCTTACCGCGGTAAACTGATTATCC 46187

QY 265 CGATCAAAATCGGCGAAATCGCGAAGTTTCAGTACCTCCGCTGATAGGCGGATCGATAA 324
DB 46188 CGCTCGTCCGAC-----CTCCATATTTACACTAGAACGATGCCAAGACTCTCTTG 46241

QY 325 CTTTGATGTTGGGAGT---CTGACCAGGATTTTGGACTACCTTAAAGTGCATCAACTGG 381
DB 46242 TTTTACCGTAGCAGAGTCTACTGAAACCGACTTCGACCACTGAAGTCTGATTCCCTCA 46301

QY 382 TAGATTCGAATTTGCGTGGCTTTTATATGTTATGTCACGGGTTATAGGACCATGC 441
DB 46302 AGATATCATGTTGTCATGCGCTCTGCCCAAGTTACTCTCTCTCAGTCTCTCCGG 46361

QY 442 AAGACTATAAGTGATCCCGCTCGTAGCGTGCAAGTAAACGTTTCTCTACCCGTTGCA 501
DB 46362 AAGGAATTCAGATGGACCAATAAATGGCTATGCAAGTCAACATCTCTCCCTGGAGTGGAA 46421

QY 502 TAGCCCTGGCTCTGACGGCACATCATTTCAATTTGCAGATGCTAAAAGTTTGTAAATGTCA 561
DB 46422 TCTGTATAGGCAACTCAGCTACACATGTTGTAGCAGATGGAGTCACTTTCAGTCAATTTCA 46481

QY 562 TCATGCTTTGGGCTATATTACAAATTTGGGAAGACCGCGGACTTGTGTCCGCGCAATC 621
DB 46482 TGAAGTATTGGATGTCATTGACCAATTCAGCGGTAAAGATCCCGCACGGTTCTTCTAC 46541

QY 622 TTCTTCCATCTTTTCGATAGATGATCAATCAAGATCTGTATGGCTAGAGGAACATTTT 681
DB 46542 CCTCTCTGCTATTACACGCTGCAGAACTATGATCAAGGACCGCAGGATGAGGCGAG 46601

QY 682 GGAACGAATGCAAGATGTTCTTGAATGTTCTCTAGATTTTGAAGCAAAACCCCTCGAT 741
DB 46602 GACATTTGGAGCGGTTTTCGAGCCAAACTCTGCAAAACACAGTTCCACATGTTACGCCGTG 46661

QY 742 TCACAAAGGTAGCAGGTACATATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 801
DB 46662 AGAACATGCTCAGAGCCACTTTTACATTTAGCGCGGAAGCAGATAGATAATCTAAAGATT 46721

QY 802 AAGTACTGAATCTCAGAGGATCGGACCCACAAATAGTCTTAACGAGTTTCAACATGACCT 861
DB 46722 GGGTTACAGGC-----AGTCTGAGAATCAATCTCTCTCTCTCTCTCTCTCTCTCTCT 46775

QY 862 GTGGATACCTATGGACATGCATGGTGGTCAAAATCA---AAGATGACGCTGATCATGAGGAAT 918

DB 46776 TAGCTTTTCATTTGGGTAAAGCTTGATCAAGACACTTGTACAAAGACACTGTAACAAAGGCCA 46835

QY 919 CATCGAACACCAAAATAGACTCGAGTACTTCACTTCACTTGTACAGCGGATTCGCCAGGACTTC 978

DB 46836 ACGAGGAAGACAAAGGATGAAGTATTTCACCTTGATGATCAATGTGGATTCGAGGAATCGCG 46895

QY 979 T-----GACGCCCGCTGCTCGCTAACTACTTTGGCAACTGTCTTGGCTCATGCGTTG 1032

DB 46896 TCAAGTACACACACCAATACACAAACATACTTTGGCAACTGTATGGCTCTCTGTTATCG 46955

QY 1033 CAAAAGACCCATTGAAAAGAGTTAGTTGGGATAAAGGCTTCTTGTTCAGTGTTCAGCTA 1092

DB 46956 TATCAGTCAAGAAACACAGATTTGCTAGGAGAAAAATGCGTTTTGGCGCTTCAGATGCAA 47015

QY 1093 TTGGAGAACCCATTGAAAAGAGTTGCAACACGAAAAAGCGCTTCTTCAGATGCAAAAA 1152

DB 47016 TCACAGAGAGAATCAAGATATGTTATCAAGCG---ATCTGTTGAAGACAGCACCAAGAT 47072

QY 1153 CTTGGTTATCGGAATCTAATGAATCCCTTCAAAAAGATTTCTCGGGATTACCGGATCGC 1212

DB 47073 GGGACAAAGGATACGTAATGGTCACTGCCATTTATCCAACCTCAATTCAGGAGCTC 47132

QY 1213 CTAAGTTCGATTCGTATGCTGTAGATTTTGGATGGGAAAGCCTGCAAAATTTGACATTA 1272

DB 47133 CGAAATTTGGGACTGTATGACATGGATTTTGGATTAGGAAAGCCCTGCAAAATGGAGATCG 47192

QY 1273 CCTCTGTTGATTTATGCAGAAATGATTTATGTTAGTTCAGTCCAGGGATTTTGAANAAGT 1332

DB 47193 TGCACATAGAAACAGGTGTTCTTCATCGCATCTCAGATCCAGAGACGCGCAAGATGGAG 47252

QY 1333 TGGAGATTCGAGTATCATTTGCTTAAGATTCATGATGATTCATTTGCAAAATCTTTTGAAG 1392

DB 47253 TTTGAGTTGGAATAGCAGTACAGAGAAAGAAATGATGTTGTTGACTCAATTCCTACAAC 47312

QY 1393 AAGG 1396
DB 47313 AAGG 47316

RESULT 9
AC002560 101176 bp DNA PLN 19-JUL-2000
LOCUS Genomic sequence for Arabidopsis thaliana BAC F21B7 from chromosome
DEFINITION I, complete sequence.
ACCESSION AC002560
VERSION AC002560.2 GI:9211167
SOURCE HTG.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 101176)
AUTHORS Khan,S., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Kim,C.,
Shinn,P., Altafi,H., Bei,Q., Chin,C., Chlou,J., Choi,E., Conn,L.,
Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T., Lam,B.,
Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N.,
Nguyen,M., Palam,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A.,
Thavert,A., Toriumi,M., Vaysberg,M., Yu,G., Federspiel,N.A.,
Theologis,A. and Ecker,J.R.
Genomic sequence for Arabidopsis thaliana BAC F21B7 from chromosome
I
Unpublished
2 (bases 1 to 101176)
Ecker,J.R.
Submitted (25-SEP-1997) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
3 (bases 1 to 101176)
Ecker,J.R.
Direct Submission
TITLE

JOURNAL Submitted (13-NOV-1997) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE 4 (bases 1 to 101176)
AUTHORS Ecker,J.R.
TITLE Direct Submission
JOURNAL Submitted (26-JAN-1998) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE 5 (bases 1 to 101176)
AUTHORS Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q.,
Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C.,
Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N.,
Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J.,
Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H.,
Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M.,
Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
TITLE Direct Submission
JOURNAL Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA
REFERENCE 6 (bases 1 to 101176)
AUTHORS Ecker,J.R.
TITLE Direct Submission
JOURNAL Submitted (15-JUL-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE 7 (bases 1 to 101176)
AUTHORS Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q.,
Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C.,
Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N.,
Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J.,
Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H.,
Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M.,
Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA
COMMENT On Jul 15, 2000 this sequence version replaced gi:2618677.
FEATURES
source location/Qualifiers
1..101176
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F21B7"
join(736..1038,1276..1713,1874..2119,2205..3032,
3117..3274,3483..3610,3725..4134,4313..4435)
/note="unknown protein; similar to ESTs dbj|AV529515.1,
dbj|AV529266.1, dbj|AV54950.1, dbj|AV523761.1,
dbj|AV523761.1, dbj|AV96703.1, and dbj|AV546743.1"
/codon_start=1
/evidence=not_experimental
/product="F21B7.1"
/protein_id="AAF86508.1"
/db_xref="GI:9280639"
/translation="MOPLPARGSDHEGFNMNHPHLLVAVAGDETNGTGLGHFSQNGSL
ARGDSKAGDAINVTPTTFRFSLRSHSVVYVLRPSVCMIRCSRRVAVGLANCI
YCDATLTLENKSVLYVPQVROCTRVNCGVGMVAPRIAYAKSSMTKTR
LSQFTTSFSLSPSSSGSSGFARVAMESSQLANGILNGDMGYKLSKVCQDML
PDGSTSPASNALKVGVSQDAENAGVAVKDLVSGALVQSFKAHTSPISALCFDP
SGTLTLVATVCGNNINVFQIMPSSHINAPDGLSYEWESHVHLFKLHRGITSIVODI
CFSQSQSWAIISSKGTCHIFVLNSSGDAAFQCEGEPTRLPASLPMWFTQSLS
NOQSLSPPTVALSVRSRIKYSFGMLNTVSNATTAATGKVFVPSGAAVAFHKSVTH
DLQNSPTNALEHLVATPSGHVQHELLPVCETSENGLRVQKTSHVQVQEDDLRV
KVPTIQWDCVCRSDMLETERLPKSITEKQYDLETVSNHLTSHEDACLSDMNSHPS
EDKYLKSCSKPERSHCYLSNFVEVKVTGMLPWNQNSKLSFHWMDSPRDSSTGGEF
EIKVPAHELEIKQLLPVDFHSTKATLEDRFSMKCYHTSATSGHQVNGYKTEKNN
INCHKPGSTSESSBEGSTKQENLHSDHMSNISKSLSLPTVNGIYKTEKNN
ANGMEKPVTKLTKRTITNGTTPPTLTDVSNQFMSLTGKPMGFGFALHEHC
KAVADPKEEHLKKKLDPTVNVHLNWNNTKLGQDKMWNQSVLNAFGVCIYKDFPL
RVGNPNFPLLLYKILRK"
complement(4838..6223)
CDS
/note="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F21B7.2"
/protein_id="AAF86541.1"
/db_xref="GI:9280672"
/translation="MASCIOQLQEN"VFSTNNRESFVLKRODALSEVLVPPYPLSGRLREVENGLK
NLDDIICARVFTPSVYCPSTNNRESFVLKRODALSEVLVPPYPLSGRLREVENGLK
EVFGREGQVLMVANSNDMDLADGLTVPNANPLFIERNPGEEAYKILEMPLLIAQ
PTVTCGGFSLGIRLCHICDGFAMOFLOKSWAATKATKGLIADPPVMDRETFKPRN
PMVKYPHHELYPIEERSNLNLSMDTKPQKCYRISKFEQCRVKSIAQEDPTLVCS
TFDMAAHWRSWVKALDYKPLDYNLRTFSVNVNTRLETKLKRGFYGNVGLACAM
SSVESLNSLSTRLVQDARLVSEDYLRSMVDYVVKRKRLEFGGLKLTITQWTR
FENYETADCGKGPVAGPIDLRPTQCVLLPQGGVESGNDQSMVVCICLPPTAVHT
FTRLNSLNDHK"
complement(join(6700..7357,7482..7884,8620..8839,
8920..9241,9758..10380))
/note="unknown protein; similar to ESTs gb|T04168.1,
gb|T21964.1, and gb|AW004462.1"
/codon_start=1
/evidence=not_experimental
/product="F21B7.3"
/protein_id="AAF86540.1"
/db_xref="GI:9280671"
/translation="MGHDSFCYLIVLRCLRCGIIALMQICALQKKERRSKMESSDRS
SOAKAFDETGTGKGLVAGSIKEIPAMFHTPDTLTSLKOTAPPQOOLTPTVDLKG
SMDLISRRSVVEKIGDAERMGFQVNVHGISVEVMERKEGIRRRHEQDPEVKRFY
SRDTRDLYISNIDLHTCNKAANWRDTLACYMAPDPPKLODLPVACGEIMMEYSKOL
MITLLELLELLSEALGNPNHLKDMGCAKSHIMFGQYPPCPQDPLTGLISKHTDFS
ITILLQNTIGGLQVIVHDCQVSPVGVALVINIGDQLLISNDKFTSAEHRVIANGS
SEPRI SMPCVSTFMKPNRIYGIKELSEQNPARYDLTITFESNTFSRLVNRITP
IPALFRATPATLASLSPPPKHLITPTVDLKASVVEKIGEAKEKWLRLNHLGYP
VEVERMIOGIRGFHEQEPAKKRFYSRDTRDLYVSNHDLONSEASNRDTLGCYT
APPPREDLPVACGEMILEYKSEIMSLGERLLELSEALGLSHHLKDKMCAKQYM
VGHYPPCPQDPLTIGINKHTDISFLTVLQDNVGGVQFVHEQVTDVTPGALVIN
IGDFLVQKTKMLLRSCFTSCIMWGLILFYCOLINDKFTISAHRVIANGESEPT
SVAIVSTFMRAYSRVYGPICKLLSAENPAKYRDCITLTFSTFSSKTLDAPLHHEK
I"
10833..11348
/note="unknown protein; similar to EST emb|Z30724.1"
/codon_start=1
/evidence=not_experimental
/product="F21B7.4"
/protein_id="AAF86509.1"
/db_xref="GI:9280640"
/translation="MRGFSVSSWLATSVIEGVFLGSTTDLDAVAEADAVDMVVS
ACLMVKNKDLRLNLLRVPVAGLSEPPFSLRFLSPSIRIEGCFAGKFSDLRFRWSE
AMKCVVVLRGWGLDQIGWRSVGNWVFSRVFVAGPSSSLRDLRODSTANDLVPPVV
LQSSGFCIRY"
join(11965..12033,12173..12303,12448..12519,12596..12685,
12756..12915,12985..13041)
/note="similar to ATHP1 dbj|BAA37110.1"
/codon_start=1
/evidence=not_experimental
/product="F21B7.5"
/protein_id="AAF86510.1"
/db_xref="GI:9280641"
/translation="MNTIVVAQLQRFODYIVSLYQOQGLDNOFSELUKLQDECTPDF
VAVSLFPDCKSLINTMSISLERPDNDVKQDVGCHOLKGSSSSLCFLVSGARR
VNVCSISFKCCDVQNRGYSNHNCDNDIWHRCRLRCLQQVDYEFYKMLTKLQDLF
VNVKFCENLVLVLLLEKQILQAGGTIPQVDIN"
16495..17682
/note="similar to protein kinase pir|A57676; similar to
ESTs gb|A1993651.1, dbj|AV538995.1, dbj|AV522534.1, and
dbj|AV538995.1"
/codon_start=1
/evidence=not_experimental
/product="F21B7.6"
/protein_id="AAF86511.1"
/db_xref="GI:9280642"
/translation="MRFTKLAVCMFLIRFGFTEALLDPVDFLALQAIKSLDLDLP
SKFFESWDFTSDFCGFAGVYCNKGDKVSLNGLDPRAGSPGLSGRIDRPAIKLSALTET
SIVPGRIMGALPATISOLKDLRFLAISRNFIISGEIPASLGEVGRGLRDLTLDUSYNOLTGT

ISPSGSLBELSNLILCHNNHLTGSIIPPELSQTLTRIDLKRNSLTGSIISPASLPSPSQ
LSLANQLTGSVYHRLNQLNLDUSLNRETTGTIPNRFAPFPIINLQJRNFFGL
IQPANAQVTSIVDLSSYNRFGSLPLSSVENVLIASNRFTEGPVASFVERLLSNIIQ
TLYLQHNFLLTGIOISPAAEIPVSSSLCLQYNMVPPIQTCPLKAGPKQKTRPTTQCTE
JBC.

complement(join(18462, .18515,18612, .18742,18840, .18975,
19064, .19139,19214, .19410,19497, .19595,19668, .19752,
19865, .20103,20179, .20253,20345, .20649,20798, .20902,
21088, .21216,21356, .21437,21513, .21608,21780, .21849,
21929, .22045,22408, .22302,22586, .22659,22810, .22888,
23041, .23101,23206, .23513,23411, .23556))

[illegible]

Query Match	6.4%	Score 109.2:	DB 12:	Length 101176;
Best Local Similarity	44.9%	Pred. No. 4.6e-17;		
Matches 589:	Conservative	0:	Mismatches 708:	Indels 15; Gaps 4;

Qy	85	CGCTACCGGTAAACATCTTTCGATATCCCTCGGTTGGCACTTGGAATAAGATGCAGTCCCTTC	144
Db	37326	CTCTCCTCTCACCTTCTTCGATCGCCCTGGCTCTCTCTCCCACTCGCCGATTCTCTCT	37385
Qy	145	TGTTTTAGGACTTTCCGTACCACAAGACACATTTCTTGGACACTGTATCCCTAATCTTA	204
Db	37386	TCCTCTCTTACCAAACTCAAATCTTCTTCCCTCCAAGACTTGTACCCAACCTCA	37445
Qy	205	AGGCTCTTTGTCTCTCACTCTAAAACACTACGTTCCTCGCTTAGCGGAAATTTGTGTATGC	264
Db	37446	AACATTCCCTCTCCATCACTCTCCAACATTTCTTCCCTTAGCGCGGTAACATGATTATCC	37505
Qy	265	CGATCAATCGGGGNAATGCCGAAGTTTCAGTACTCCCGTGATGAGGGGACATCGATAA	324
Db	37506	CGCCTCGTCC---CGACCCCTCATATTACACTACAAGCTGGCGAAGACTCTTGTGTTT	37562
Qy	325	CTTTGATCGTTGCGGAGTCTGACCAAGATTTTGACTACTCTTAAAGGTCATCAACTGSGTAG	384
Db	37563	TCACCGTAGCAGATCTACTCAAAACCGACTTCGACCAACTCAAGTCTGATTCCCTCAAAG	37622
Qy	385	ATTCCAATGATTTGCATGGCCCTTTTTTATGTTATGCCACGGGTTATPAGGACCATGCAAG	444
Db	37623	ATATCAGTGTTCGATGGGCTCTGTGCCAAGTTTACCTCTCCCTCAGCTCTCTCCGGAAG	37682
Qy	445	ACTATAAGTGATCCCGCTCTAGCCGTGCAAGTAAACGGTTTTTTCCTAACCGTGGCATAG	504
Db	37683	GAA TTCAGATCGACCAATATATGCAATGCAAGTCACCATCTCCCTCGGAGCTGGAATCT	37742
Qy	505	CCGTGGCTCTGACGGCACATCATTAATGCAAGATGCTAAAAGTTTGTAAATGTCATCA	564
Db	37743	GTATAGGCAACTCAGCTACACATGTGTAGCAGATGGAGTCACCTTCAGTCAATTTTCATGA	37802
Qy	565	ATGCTTGGGCTATATTAACAAATTTGSGAAAGACGGGACTTCTTGTCCCGGAATCTTC	624
Db	37803	AGTATTTGATGTCATTTGACCAANATCCAGCGCAAGATCCGGCCACGGTTCTTCTACCTT	37862
Qy	625	TTCCATCTTTTCGATAGATCGATAATCAAGAATCTGTATGGCTTAGGGAACAACTTTTGA	684
Db	37863	CTCTGCCTATTTCAGCTGCAAGAACATAATCAAGGACCCAGGAGAGGTAGCCCGAGGAC	37922
Qy	685	ACGAANTGCAAGATGTTCTTCAATATGTTCTAGATTTGGAAGCAAAACCCCTCGATTCA	744
Db	37923	ATTTGGAGCGGTTTTGGAGCCAAACTCTGCANAACACAGTTTCACATGTTACGGCTCGAGA	37982
Qy	745	ACAAAGGTACGAGCTACATATATCTCTCTCCCTTGCTGAAATCCAGAAGCTAAAGAACAAAG	804

Db	37983	ACATGGTCAGAGCCACCTTTTACATTTGAGCCGGAAGCAGATAGCAATCTCAAGAGTTGGG	38042
Qy	805	TACTGAATCTCAGAGGATCCGAAACCGCAATACGCTGTAAACGACGCTTCAATACGCTGTG	864
Db	38043	TTACAGAGC-----AGTCACGAATCAATCTCCTGTTTCTACCTTCGTGTGACTCTAG	38096
Qy	865	GATACGTTATGCACATGCATGTCCTCAATCAAAAGATGAGCTCGTATCAGAGGAATCATCGA	924
Db	38097	CTTTTCATTTGGGTAAAGCTTGATTAAGACACTGTGTTCAAGACAGTGAACAGAGGCCAAGG	38156
Qy	925	ACGACGAAATGAGCTCGAGTACTTTCAGTTTACAGCGGATTTGCCGAGGACTTCTGAGCG	984
Db	38157	ACGAAGCTTTTCACTTTGATGATCAACGTGGATTCAGGAAATCGTCTCAAGT---ACACAC	38213
Qy	985	CCCGGTGTCGCCCTAACTTATTTGGCACTGTCTTGGCTCATGGTTGCCAAAGCAACAC	1044
Db	38214	AACCAATACCACAACATACCTTTGGCACTGTATGGCTCTCGGTATCGTATCAGTCAAGA	38273
Qy	1045	ATAAAGAGTTAGTTGGGGATAAAGGGCTTCCTTTGTGCAAGTTGCAGCTATTGGAGAAAGCA	1104
Db	38274	AACAGATTGCTAGGAGAAAAATCGTCATGGCTTCAGATGCAATCACACGGAGAA	38333
Qy	1105	TGAAAAAGAGTTGCAACAACGAAAGCGGTTCTTGAGATGCAAAAACCTTTGGTTATCGG	1164
Db	38334	TCAAAGATATGTTATCAAGCG---ATCTGTTGAAGACAGCACCAAGATGGGGACAAGGAG	38390
Qy	1165	AATCTAATGGAAATCCCTTCAAAAAGATTCTCGGGATTACCGGATTCGGCTTAAGTTTCGATT	1224
Db	38391	TACGTTAATGGGTCAATGTCCTCCATTTATCCAACTCAATTCGAGGACCTCCGAATTTGGAC	38450
Qy	1225	CGTATGGTGTAGATTTTGGATGGGGAAGCGTCAAAAATTTGACATTTACCTCTCTGTTGATT	1284
Db	38451	TGTATGACATGGATTTTGGATTAGGAAAGCCCTGCAAAATGGAGATAGTCACATAGAAA	38510
Qy	1285	ATGCAGAAATTCGATTTATCTGATTCTAGTCCAGGGATTTTGA AAAAGGCTGCAGATTTGAG	1344
Db	38511	CAGTGGTTCTATCGCATTTCTCAGAGTCCAGACGGCAGCAATGGAGTTGCGATTGGAA	38570
Qy	1345	TATCATTCGCCCTAAGATTTCATATGGATGCATTTGGCAAAAATCTTTGAAGAAGG	1396
Db	38571	TAGCACTAGAGAAGAAGAAATGGATGTTGTTGACTCTACTTCTACAAAGG	38622

RESULT	10
LOCUS	AX025514
DEFINITION	Sequence 40 from Patent WO0032789.
ACCESSION	AX025514
VERSION	AX025514.1 GI:10187184
KEYWORDS	.
SOURCE	Citrus limon.
ORGANISM	Citrus limon
	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eucotyledons; core eudicots; Rosidae; eurosids II; Sapindales; Rutaceae; Citrus.

REFERENCE
AUTHORS
1 (bases 1 to 1648)
Tunen,A.J., Verhoeven,H.A., Luecker,J., O'Connell,A.P. and Van
Tunen,A.J.
TITLE
Fruit flavour related genes and use thereof
Patent: WO 002789-A 40 08-JUN-2000;
AHARONI ASAPH (IL); VERHOEVEN HARRIE ADRIANUS (NL); LUECKER JOOST
(NL); CPRO DLO (NL); CONNELL ANN PATRICIA O (NL); TUNEN ARJEN

FEATURES	source	Location/Qualifiers
		1..1648
		/organism="Citrus limon"
		/db_xref="taxon:2708"
CDS		52..1527
		/note="CDNA
		Lemon acyl transferase"
		/codon_start=1
		/protein_id="CAC090866.1"

```
/db_xref="GI:10187185"
/translaton="MAAIENRVLTKHEVTKVFPVNPNSKTSFTLIDLTYDFWFVK
NPPVERLFFVEMDILTWDLNSELPLKHLSTLLHYLPLAGHIMWLDAAKPAVY
YFPDNGVSEFAVEMSECHAGFHLSGNGIRQAVEFHLVPOLSLTDKAEVIAIQ
ITLFPNPGESIGVSSHAIIDGKSTSLFLKSWAYLCKQLCHLHPCLSPELTLLDRT
VINDPTGDMQLNKWVWSDNPDQKIRSLKLPFLDSESLNKLVRATELITREDIT
KLRRHWHQJLSKSKSKQVRLSTFVLTLAYFVCMKAKLAKAKTAEAAAGNDEIKN
IIVGTADYRSRDLPPPLNFGNCRHGETAKASDFVQENGVAFAEMLSDMVKGI
DAIAEANDKVEISILILEKGAIFSVAGSTQFDVYGSDFGWRPKKVEISVDRTO
AISLABRDRDGGGVGVGVLEKQOMEVFEVSFADGLKNDLV"
BASE COUNT 464 a 344 c 354 g 486 t
ORIGIN

Query Match 4.8%; Score 81; DB 9; Length 1648;
Best Local Similarity 46.6%; Pred. No. 4.9e-10;
Matches 643; Conservative 0; Mismatches 660; Indels 76; Gaps 9;

Qy 80 GTTATFCGTACCGGTAAACATTCTTCGATATCCCTCGTGTGCACTTGAATAAGATGCAGTC 139
Db 141 GTTATCTCGATCTCACTATTTCGACTTTTCTGCTTCAAGAATCTCTCTGTGGAACG 200
Qy 140 CCTCTCTGTTTACGACTTTCGTAACCAAGAACACATTTCTTGACACTG---TTATCCC 196
Db 201 CCTCTCTCTATGAGATGACTGACTTGACGTGGGATTTATTTCAACTCGGAGATCCTCCC 260
Qy 197 TAATCTTAAGGCTCTTTGCTCTCACCTCTAAACACATACCTTCCGCTTACGGGAATTT 256
Db 261 AAGCTGAAGCACTCCCTTCTTCACTCTCTTCACTTCACTTACCTCTCTCTGTCAGT 320
Qy 257 GTTGATGCCGATCAATCGGCGAATCGGAAGTTTCAGTACCTCCGCTGATGAGGCGA 316
Db 321 CATGTGGCGCTGGA---TGCGCGAAGCCTGCCCTCTACTACTTTCCCGCAACAAACGA 377
Qy 317 CTCGATACTTTGATCGTTCGGGAGCTGACAGGATTTTGACTACCTTAAGGTCATCA 376
Db 378 CGGCTTTTCATTCGCGATGCTGCTGAGT-----GTCCTTCGAGTGCCACGAGGCTTCCA 431
Qy 377 ACTGATAGATTCGAATGATTCGATGGCTTTTATGTTATGTCACGCGTTATAAGGAC 436
Db 432 TCACCTCTCCGCAACGGAATCCGCAAGCAGTTGAATTTTCATCTCTGTGCCCCAGTT 491
Qy 437 CATGCAAGACTATAAAGTATCCCGCTGCTGAGCGTGCAGAGTACCGTTTTCCTAACCG 496
Db 492 GTCGCTTACGACGACATAAAGCTGAGGTAATTCGCAATCAACATAACACTGTTTCGGAATCA 551
Qy 497 TGGATAGCCGCTGCTGACGGCACATCTCAATTCGACATGCTAAAGCTTTTGTAAAT 556
Db 552 AGGCTTCTCAATGCTGTTTTCATCTCACCATGCAATCTTGATGGAANAATCTCGACCTT 611
Qy 557 GTTCATCAATGCTTGGGCTATATTAACAAATTTGGGAAG-----ACGCGGACTTGT 610
Db 612 GTTCTGAAATCTTGGGCTTATTTGTGCAACAATTAACAATATGCCATCACCCTTGTTT 671
Qy 611 GTCGCGAATCTTCTTCATCTTTGATAGATGATGATAAAGATC----- 657
Db 672 GTCACCTGAACCTCCCTCTTCGACCGGACTGCTCAAAAGATCGACGAGGTCAGGA 731
Qy 658 -----TGTATGGCTAGAGGAACATTTTGAAGCAATGGAAGCAATGCAAGAT 698
Db 732 CATGCTGCAATGAATAAGTGGTGTGGGTCGGATATTTTCGATCCCGAAGATACG 791
Qy 699 GTTCTTGAATGTTCTCTAGATTTTGGGAAGCAACCCCTCGATTCACAAAGGTACGAGCT 758
Db 792 GAGCTTGAAGGTTTACCATTTCTAGACTCTGAGTCTCTGAA---CAATTTGTCGAGGCC 849
Qy 759 ACATATGCTCTCTCCCTTGTGTAATCCAGAGCTTAAGAAAGTAAAGTACTGAATCTCAGA 818
Db 850 ACATTTGAGTTGACGCGTGAAGATATTACGAACACTCAGGCACAAAGGTTAATCATCAGTTA 909
Qy 819 GGATCCGNAAC-----CACAAATAGCTTAAGGAGGTTCAATGACGCTGTGGATAC 869
Db 910 TCAAAATCATCAAAATCAAGCAAGTTGCTTTATCAACTTTTGTGCTTACATTAGCTTAT 969
```

```
Qy 870 GTATGGACATGCATGGTCAAAATCAAAAGATGAC-----GTCGTATCAGAGGAA 917
Db 970 GTGTTTGTTCATGGCTAAAGCTAAATTACCAAAAGCCAAACTGAAGCTGAAGCTGCA 1029
Qy 918 TCATCGAACGACGAAATGAGCTCGAGTACTTCACTAGTTTACAGCGGATGCCGAGGACTT 977
Db 1030 GCAGTATGATGAATATAAAATATATTATGTTGGGATTTCACTGGGATTTATAGGACCGT 1089
Qy 978 CTGACGCCCGCTGCTGCCCTTAACCTACTTTGSCAACTGCTTTGCGTCATGCTGTTGCAAAA 1037
Db 1090 TTGGATCTCCAATTCACCTTAATTTTGGTAACCAATGGGAGACATTTGAGACT 1149
Qy 1038 GCACACATAAAGAGTTAGTTGGGATAAAGGGCTTCTTGTTCAGTTGCACTTGCAGTCTTGA 1097
Db 1150 GCAAAAGCAAGTATTCGTTTCAAGAAATGSGGTGCTTTT---GTTGACAGATGTTA 1206
Qy 1098 GAAGCCATTGAAAAGAGTTGCAACAGCAAAAGCGGCTTTCGAGATGCAAAACCTTGG 1157
Db 1207 AGTGATATGTTCAAGAGGATCGATCGGATGCCATTTGAAGCCAATGATGATAAGGTTTCA 1266
Qy 1158 TTATCGGAATCTAATGGAATCCCTTCAAAAGATTTCTCGGGATTACGGATCGCCTAAG 1217
Db 1267 GAATATTTGGAATTTCTGAAAGAGGAGCAATGATTTTCTGTGGCTGGCTCGACCCAA 1326
Qy 1218 TTCGATTCGTATGTTAGATTTTGGGATGGGAAAGCCTGCAAAATTTGACATTTACCTCT 1277
Db 1327 TTTGATTTTACGGTTCGATTTTCGGTGGGGAGGCCCAAGAGGTTGAGATTTGTGCA 1386
Qy 1278 GTTGATATGCAAGATTTGATTTATGATTCAGTCCAGGAGTTTGTAAAGGTTGGAG 1337
Db 1387 ATAGATAGACACAAGCCTCTCTTTGGCAGAGAAGAGATGGAGGAGCGGCTTGAG 1446
Qy 1338 ATTGAGTATCATTTGCTTAAAGATTCATATGATGATCTTTCGAAATCTTTTGAAGAAG 1396
Db 1447 GTTGAGTGTGTTTAGAGACCAAAATGAGGTTTGTGAATCTGTATTTGCTGATGG 1505

RESULT 11.
AP000606/c 82360 bp DNA PLN 27-DEC-2000
LOCUS Arabidopsis thaliana genomic DNA, chromosome 3, pl clone: MTO24.
DEFINITION AP000606.1 BA000014
ACCESSION AP000606.1 GI:6045161
VERSION
KEYWORDS
SOURCE
Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui pl
clone:MTO24.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (sites)
AUTHORS Kaneko, T., Katoh, T., Sato, S., Nakamura, A., Asamizu, E. and Tabata, S.
TITLE Structural analysis of Arabidopsis thaliana chromosome 3. II.
SEQUENCE FEATURES OF THE 4,251,695 bp REGIONS COVERED BY 90 PL, TAC
AND BAC CLONES
DNA RES. 7 (3), 217-221 (2000)
JOURNAL 20363099
MEDLINE 2 (bases 1 to 82360)
REFERENCE Kaneko, T., Kato, T., Sato, S., Nakamura, Y., Asamizu, E. and Tabata, S.
AUTHORS Direct Submission
TITLE Submitted (13-OCT-1999) Yasukazu Nakamura, Kazusa DNA Research
JOURNAL Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
Tel: 81-438-52-3935, Fax: 81-438-52-3934)
ADDRESS FOR CORRESPONDENCE: kaos@kazusa.or.jp
COMMENT For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agd-graph.cgi?c=MTO24
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
```

<http://compbio.ornl.gov/Graill-1.3/>),
 GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>),
 NetGene2 (S. M. Hebsgaard, et al., CBS Technical University of
 Denmark, <http://www.cbs.dtu.dk/services/Netgene2/>) and
 SplicePredictor (Volker Brendel, Stanford University,
<http://gremli1.zool.iastate.edu/cgi-bin/sp.cgi>).
 Genes encoding tRNAs are predicted by tRNAscan-SE
 (Sean Eddy, Washington University School of Medicine, St. Louis,
<http://genome.wustl.edu/eddy/tRNAscan-SE/>).
 This sequence may not be the entire insert of this clone. It may be
 shorter because we remove overlaps between neighboring submissions.
 The 5' clone is MWE13 and the 3' clone is T13J10.
 Location/Qualifiers
 1. .82360
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone_lib="Mitsui P1"
 /clone_id="MTO24"
 5793..8151
 /notes="gene_id:MTO24.1"
 /codon_start=1
 /evidence="not_experimental"
 /product="retroelement pol polyprotein-like"
 complement(join(9270..9952,10067..10153))
 /notes="contains similarity to copia-like retroelement pol
 polyprotein"
 gene_id:MTO24.2"
 /codon_start=1
 /pseudo
 /evidence="not_experimental"
 /product="retroelement pol polyprotein-like"
 complement(join(10964..11326,11419..11702,11802..11968,
 12054..12160))
 /notes="emb|CAB68159.1
 gene_id:MTO24.3
 similar to unknown protein"
 /codon_start=1
 /evidence="not_experimental"
 /protein_id="BAB01190.1"
 /db_xref="GI:11994161"
 /translation="MGNDNKFVWIKNVPTLNSDMFSNFYVIGGCSWRVVAHSKE
 NFKESLUTLVAEDSAOKMGWSYAKIIFTLVNOISEILSORIETPMFQKSSVF
 SSETFWISCKFDEHAGFVNGEIKIVFEELEIDKLVLKSKSNPPFKTKLNNGEV
 SKDLREVPIWESIVVNGFHLVPSQVEVVKRIFEPKPDVAKREFRPTNRIVTATNV
 LUSLETLQSPREISKNOLDGACGLLRSMKAGFKLDWLKKLNEVLEKKEEESYE
 TRMRETEEMKDLKALDVGAPLRDDVV"
 complement(12716..13195)
 /notes="emb|CAB38296.1
 gene_id:MTO24.4
 similar to unknown protein"
 /codon_start=1
 /pseudo
 /evidence="not_experimental"
 complement(15206..16555)
 /notes="gene_id:MTO24.5"
 /codon_start=1
 /evidence="not_experimental"
 /product="anthocyanin 5-aromatic acyltransferase-like
 protein"
 protein_id="BAB01191.1"
 /db_xref="GI:11994162"
 /translation="WVNFNSAVNILEVVOVSPSSNSLTPLTYFDLGLKLPVDRV
 LYHVPFLTRSSLISKLSLSATLLHLYPLAGRLWDSIKTKPSIYISPPDKDAVIL
 TVAESNGDLSHLSGDEPRPATEFHSVLPELPSDESARLAVQVTFPFGQFSLGTA
 HNAVLDTKTFAMFLKWAHNCKQEALPHDLVPSLDRIYQDPTGLETKLLNRWISA
 SNKPSLKLFPSSKIIGSDILRVYTLTREDIKKLRVETESHAKOLRSTFTVITVAY
 VITCMVKMGSGDPTREVCVGFASDPESRLNPLPTFFGNCIVGSGDFDVKAEPLIEE
 GEGKGFITAVETLTGHWNGLCPENIEKNMLLPFAFKRPEGRQMSVAGSTRGIYG
 SDFMGKPKVVEIITDKDASVLSGSGGVGVCLKKDDKDDERFSLFSIGLE"
 join(21572..21604,21666..22063,22148..22642,22710..22929)
 /notes="contains similarity to Mutator-like transposase"

gene_id:MTO24.6"
 /codon_start=1
 /evidence="not_experimental"
 /protein_id="BAB01192.1"
 /db_xref="GI:11994163"
 /translation="MMRMSVSIYLQELSRVENVVKNSSIGKNYADLGSNEDMDRYATI
 VNGELGQCQEONEVIEVDVDDQDPEIGSDRDRDDGDVEYVPPHAYVSSQFKREW
 EDSISLTLEEFPSRAALHEVVDKGAFAFSGVVIKKSKERRYKNMHTCTRTSKSNW
 RMRKRGTHYELVAALLHDTPFGLLETPAPKVIKMLVOTKLGVIKYSTALRGERQAVC
 DUKGAEDSYKDIENCYLMKKNEGVTYILKLDSEKGFQYLFIALGASIEFGQANRRK
 VILIDATHIKKVSMSYRSRVRFHREYRNKNSRTIYLVDRSQMDIPDDIKALKVLPL
 PRKKKGRTKVLRFSTGERPRKQRTQNRRLRQSLQWLFLGTHSI"
 complement(join(23836..24188,24334..24547,24687..24774,
 24898..24962))
 /note="emb|CAB61970.1
 gene_id:MTO24.7
 strong similarity to unknown protein"
 /codon_start=1
 /evidence="not_experimental"
 /protein_id="BAB01193.1"
 /db_xref="GI:11994164"
 /translation="MVEETKILSRKNKXQAEPASTDAGWSGQUTLELVYGSRRV
 KRCLQGHFFAPYIREDAQGLMDHLTTSRHHTFFLKVSQHSIKLLVKGVRGLEYNST
 ELMRMLPPGPTVWVSVEGDESDSGMDREKCSRKHKDCIRTHPIAHGNCNLGTGR
 KHQEWVRNVIIHHIGSISRLPSCYILLSHGETDRRSVRPNGSDMNDQRESRPWEPMDKT
 DVGHASKPRADGRRISWASLR"
 join(25891..26085,26155..26349,26412..26534)
 /note="contains similarity to Mutator-like transposase"
 gene_id:MTO24.8"
 /codon_start=1
 /evidence="not_experimental"
 /protein_id="BAB01194.1"
 /db_xref="GI:11994165"
 /translation="MDRICESDPWDEMSAKRIMOQLEEVAMMEIPIICPCGGRIL
 DIISEKDGKGRKYECTDYKNDGLHIQKMDKAMVEVNRURREQVDNHHQIKQSLEY
 SNQEVLSFEDEIQKMGTLWRVNVWVCYRYQVFHCILSIPLYSOFSLVFSVLSLNL
 SPFLXSLYFR"
 complement(join(26878..27120,27210..27491,27693..27854,
 27950..28175,28419..28702,29026..29487))
 /note="emb|CAB77996.1
 gene_id:MTO24.9
 strong similarity to unknown protein"
 /codon_start=1
 /evidence="not_experimental"
 /protein_id="BAB01195.1"
 /db_xref="GI:11994166"
 /translation="MAPRTIPLKKRKRKRVAGTRPIDVVTKATTEPPTTTEPSATE
 QNPVATEGVEEPIPTIITVVEGEQDNNSEENSEKDENESEEESEEESEEE
 EKEEEKEEGNVGTGESSNDSTTLGEYSDDKNMDETAVENQVEIPASMEIDEET
 KAILPLSMYFPPSEYVKIKLSIRCYIHEVLVTFDKLEPEMSKSEFQNHPSFQHI
 FHMPRDPNRLMGWMLLFTARIERKEACIIAPIKTGDKAPQVDFCLKAVSOLT
 FCENFQNGRYFFDYMGLTISHTVNHGVSVTNNEKYVNSVLGCLPMELLAEAIPL
 RETFMEDIAAGDEGCPCKVRKKNHLAGFPLDTIYAEIGTQRKESRLNRPDEIY
 QDDIQALQITAGPKTAGAEDIPAVGQELPSLVSVVEMENLDKRWQDQNLNLTWIS
 DLDKRVSELEAFDEQAEKNEEMDEQDKGDPTEKDGQVEKDGDEKDGDEGVPKED
 QEKDGVPKEKPKKEKPKEDGKPGKPEANFTRIRTRTKN"
 join(33399..33608,33762..33830,34025..34098)
 /note="gene_id:MTO24.10
 unknown protein"
 /codon_start=1
 /pseudo
 /evidence="not_experimental"
 join(34168..34242,34530..35934,36128..39479)
 /note="gene_id:MTO24.11"
 /codon_start=1
 /pseudo
 /evidence="not_experimental"
 /product="retroelement pol polyprotein-like"
 39663..40357
 /note="gene_id:MTO24.12
 unknown protein"
 /codon_start=1
 /pseudo

[illegible]

Qy	898	ATGAGCTCGTATACAGAGGAATCATCGAACGACGAGAAATAGAGCTCGAGTACTTCAGTTTTTA	957
Db	15699	GAG-----GTGACCGACCGCGTTTGTCGCGTGGTTTTG	15664
Qy	958	CAGCGGATTGCGGAGGACTTCTGACGCCCGCGTGTCCGCCCTAACTACTCTTTGGCAACTGTC	1017
Db	15663	CATCAGACTTTCCGGTACAGGTTAAACCCCTCGTTACCGCCGACGPTCTTCGGGAACACTGCA	15604
Qy	1018	TTG-----CGTCATCGGTTGCAAAAGCAACACATAAAGAGTTAGTTGGGGATAA	1066
Db	15603	TTGTCGGGTCCGGAGATTTTCACGCTGAAGCAGAACCAATATTGGAAGAAGAGNAGNA	15544
Qy	1067	AGGGCTTCTGTTGCGAGTTGCGAGCTATTGGAGAGGCCATTGAAAAGAGGTTGCCACAACGA	1126
Db	15543	AAGGTTTCATCACGGCGGTGGAGACTTTGACCGGATGGTCAACGGATTGTGTCGGGAGA	15484
Qy	1127	AAAGGCGTTCTTCGAGATGCACAAAACCTTGGTTATCCGAATCTAATGGAATCCCTTCAAA	1186
Db	15483	ATATAGAGAAATATGCTTTTACCGTTTGAAGCGTTTAAAGAGGATGGAACCGGGAAGGC	15424
Qy	1187	AAGATTCTCCGGGATTACCGGATGCCTTAAGTTCGATTTCGATCGTATGGTGTAGATTTTGGATG	1246
Db	15423	AAATGATATCG--TTGCTGGTCGACCGCGTTAGGGATTTACGGGCTGATTTTGGGTG	15366
Qy	1247	GGGAAGCCTCGAAAATTTTGACATTACCTCTGTGTGATTATCGAGAAATGATTTATGTGAT	1306
Db	15365	GGGAAAACCGGTCAAGGTAGAGATTGTGACGATTGATAAGGATGCTTCGGTTTCATTGTC	15306
Qy	1307	TCAGTCCAGGATTTTGA AAAAGGTGTCGAGATTCGAGCTATCATTTGCCTAGATCATAT	1366
Db	15305	GGAGTGGAGATGGAAGTGGTGGTGTGTGAGTTGGTGTGTTGTTTGAAGAAGATGATGT	15246
Qy	1367	GGATGCATTTGCCAAAAATCTTTGAAGAAGGCTTTTGTCTTTGTGCATAGTCTCTTTAATA	1426
Db	15245	GGACGTTTTTGGTTCATTATTCTCCATCGGTTGGAATAAAAAAGAGAACGTTTTGGTTTC	15186
Qy	1427	GAACCATATTTGCTGCGCAATAAGTACCAAGTCCCTTTAGTAACACTAC	1473
Db	15185	TAGTTTTATGTTGTTTTTAAATCCACATCATCTTTAATTATTAC	15139
RESULT 12			
AC074226	AC074226	341064 bp	DNA
LOCUS	Arabidopsis thaliana chromosome 1	clone IGF-F27B9,	*** SEQUENCING
DEFINITION	IN PROGRESS *** 5 unordered pieces.		
		HTG	15-SEP-2000

RESULT	12
AC074226	
LOCUS	341064 bp DNA HTG 15-SEP-2000
DEFINITION	Arabidopsis thaliana chromosome 1 clone IGF-F27B9, *** SEQUENCING IN PROGRESS ***, 5 unordered pieces.
ACCESSION	AC074226
VERSION	AC074226.3 GI:10140624
KEYWORDS	HTG; HTGS_PHASE1.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
AUTHORS	Lin,X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T., Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
TITLE	'Arabidopsis thaliana' IGF' BAC F27B9' genomic sequence near marker ,,
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 341064)
AUTHORS	Town,C.D. and Kaul,S.
TITLE	Direct Submission
JOURNAL	Submitted (20-JUL-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, cdtown@igr.org
COMMENT	On Sep 15, 2000 this sequence version replaced gi:10086523. * NOTE: This is a 'working draft' sequence. It currently * consists of 5 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown.

- * This record will be updated with the finished sequence
- * as soon as it is available and the accession number will
- * be preserved.

*	14021	contig of 14022 bp in length
*	14022	gap of unknown length
*	14023	contig of 21041 bp in length
*	14073	contig of unknown length
*	35113	gap of unknown length
*	35114	contig of 51695 bp in length
*	35164	gap of unknown length
*	86858	contig of 83628 bp in length
*	86859	gap of unknown length
*	86909	contig of 170586 bp in length
*	170537	contig of 170478 bp in length
*	170587	gap of unknown length
*	341064	contig of 170478 bp in length

FEATURES

SOURCE

```

source
1. .341064
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="IGF-F27B9"
BASE COUNT 110402 a 59862 c 59862 g 110479 t 459 others
ORIGIN

```

Query Match 4.2%; Score 71.8; DB 75; Length 341064;
Best Local Similarity 46.2%; Pred. NO. 1.9e-07;
Matches 604; Conservative 0; Mismatches 627; Indels 76; Gaps

[illegible]

QY 427 TTATAGGACCATGCAAGACTATAAAGTGATCCGCTCGTAGCCCGTGCAAGTAACCGTTT 486

Db 65389 -----GGTTCCGACGAATCAGCTCGGTTCTCGCCGTTCAAGGTACGTTTT 65435

[illegible]

Qy 547 GTTTTGTAAATGTTTCATCAATGCTTGGCGCTATATTAAACAATTTTGGGAAGACGCGGACT 606

Db 65496 CAACAGCCATGTTTCTCAAGCATGGG-----TCACAACGTGAACAAGAACACAAG 65549

Qy	607	TGTTGTCGCGGAATCTTCTTCCATCTTTTCGATAGATCGATAATCAAAGATCTGTATGGCC	666
Db	6550	CTCTGCCACACGAGTCTAGTACCGTCTTTAGATCGTATCTATTCCTTCAGACCCCAACGGGAC	65609

QY	667	TAGAGGAAACATTTTGGACGAAATGCAAGATGTTCTTGAATGTTCTCTAGATTGGAA	726
----	-----	---	-----

Db 65610 TCGAGACAAAACCTCTTGAACCGGTGGATATCAGCTTCCAACAACAACCAAGCTTAAAC 65669

QY	727	- - - - - GCAACCCCTCGATTTCACAAGGTACGAGCTACATATGTCCTCCTCCTTG	777
Dh	65670	TGCCTTCTCTCTTTCTCCAAAATAAATTGTCGAATGCCCATATATCTCTCCATACCTAACGGCCAACTACCTCACCGCTTT	65720

QY 778 CTGAAATCCAGAAAGCTAAAGAAACAAGTACTGAATCTCAGAGGATCCGACCGACAAATAC 837

Db 65730 AAGACATCAAGAACCTA---CGTGACGAGTGGAGACCGAGTCAACGCTAAGCAGCTCC 65786

QY 838 GTGTACGACGTTACAAATGACGTGTGGATACGTATGGACATGCATGGTCAAAATCAAAAG 897

D0 95787 9ATTGTCGACATTTTGTGCATTAACGCTTACGTTATAAACAATGCAATGGTGAAAGATGCGGTC 65884

Qy	898	ATGACGTCGTTATCAGAGGAATCATCGAACGACGAAATATGAGCTCGAGTACTTCAGTTTTTA	957
Db	65847	GAG-----GTGACCCGACCCGGTTTGTTGTCGCTTGTTTTTG	65882
Qy	958	CAGCGGATTCGCGAGGACATTCGTAGCCGCCCGTGTGCGCCCTAACTACTACTTTTGGCAACTCTGC	1017
Db	65883	CATCAGACTTTTCGGTTCAGCGTTAAACCTCCGTTACCCGCCGACGTTCTTCGGGAACTGCA	65942
Qy	1018	TTG-----CGTCNTCGGTTGCAAAAGCAACACATAAAGAGTTAGTTGGGGATAA	1066
Db	65943	TTGTCGGGTCGGAGATTTCGACGTGAAGCAGAACCAATATTGGAAGAGGAGAAGAA	66002
Qy	1067	AGGGCTTCTGTTGCGAGTTGCGAGCTATTGGAGAGGCCATTGAAAAGAGGTTGCGACAACGA	1126
Db	66003	AAGGTTTTCATCACGCGGTGGAGACTTTTGACCGGATGGCTCAACGGATGTGTCTCGGAGA	66062
Qy	1127	AAAGGCGTCTTTCAGATGCAAAACATTTGGTTATCGGAATCTAATGCAATCCCTTCAAA	1186
Db	66063	ATATAGAAAGAAATATGCTTTTACCGTTTGAAGCGTTTAAAGAGGATGGAACCGGGAAGGC	66122
Qy	1187	AAGATTTCTCGGGATTACCGGATCGCTTAAGTTCGATTTCGATGGTGTAGATTTTTGGATG	1246
Db	66123	AAATGATATCG--TTGCTGGGTGACCCGGTTAGGGATTTACGGGCTGATTTTGGGTG	66180
Qy	1247	GGGAAGCTGC AAAATTTGACATTTACCTCTCTTGATTATGCAGAAATTTGATTTATGTGAT	1306
Db	66181	GGGAAACCGGTCAAGGTAGAGATTGTGACGATTGATAAGGATGCTTCGGTTTTCATTCTGC	66240
Qy	1307	TCAGTCCAGGATTTTGA AAAAGGTGTGGAGATTTGGAGATATCATTCCTTAAGATTCATAT	1366
Db	66241	GGAGAGTGGAGATGGAAGTGTGGTGTGTTGAGGTTTGGTCTTTGTTTGAAGAAGATGATGT	66300
Qy	1367	GGATCATTTGC AAAAATCTTTTGAAGAGGCTTTTGTGCTTTTGTGCATAGTCTCTTTAATA	1426
Db	66301	GGAAGCTTTTGGTTCATTAATCTCCATCGGTTTGGAAATAAAAATGAGAACGTTTGTGTTTC	66360
Qy	1427	GAACCATATTTGCTGCAATHAAGTACCAAGTCCCTTTAGTAACTAC	1473
Db	66361	TAGTTTTATGTGCTGTTTTTAAATCCACATCATCTTAAATTCATTAC	66407

RESULT 13
AB016892

LOCUS	ABO16892	56237 bp	DNA	PLN	27-DEC-2000
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, pl clone:MXF12.				
ACCESSION	ABO16892	BA000015			
VERSION	ABO16892.1	GI:3449333			
KEYWORDS	Arabidopsis thaliana (strain:Columbia) DNA, clone lib:Mitsui pl				

Source	Organism
Macedonia (Greece)	Arabidopsis thaliana clone:MXF12.
Taiwan	Arabidopsis thaliana clone:TAIR106

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS
1 (sites)
Asamizu, E., Sato, S., Kaneko, T., Nakamura, Y., Kotani, H., Miyajima
and Mabuchi, S.

TITLE: Structural analysis of *Arabidopsis thaliana* chromosome 5. VIII. Sequence features of the regions of 1,081,958 bp covered by

JOURNAL
MEDLINE
DNA Res. 5 (6), 373-391 (1998)
seventeen physically assigned P1 and TAC clones

REFERENCE
2 (bases 1 to 66237)
Nakamura, Y.
Nucleic Acid Res.
Direct Substitution

Journal of Plant Gene Research
Submitted (18-AUG-1998) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Direct Submission

COMMENT

Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
Tel: 81-438-52-3935, Fax: 81-438-52-3934)
Address for correspondence: kaos@kazusa.or.jp

For the latest information on annotation of this clone, please send requests for correspondence: kazu@kazusa.or.jp
http://www.kazusa.or.jp/kaos/cgi-bin/aggd_graph.cgi?c=MXF12

Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <http://compro.ornl.gov/grail-1.3/>), GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi>). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K15E6 and the 3' clone is K3K3.

FEATURES

Location/Qualifiers

1. 66237

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db_xref="taxon:3702"

/chromosome="5"

/clone_lib="Mitsui P1"

2613. 5234

/note="gene_id:MXF12.2"

/codon_start=1

/evidence=not_experimental

/product="receptor-like protein kinase"

/protein_id="BAB10824.1"

/db_xref="GI:10177545"

/translation="MIRHALLIPSILVSTPIVGEATSTVEPTDVFLENGDPTSNVND
VSGRNTAENQKILSNLUNASTAQASQESGVSIPTATIRFSEPTSPPTPG
SNFLRIFYTRIGSGFNAVKSEFVSKNGFTLNNFSAADLTAKSKQTEFIKEFI
IPVQYLNTFTPSLDSAPFNGEIVSIPNRFYSGGDDVDITNIGSSVDFHIENST
AFETVRLNLTGVDGSGFRWSDDEILSESGISIPVDIKINTKPTSYVA
PDVYATSRMGNAHPQNLNLTWLTVDAGFSLVRLHFCETLSEVNGKQRFV
SFIENQATLFMGSGSWIPMWLDVTYIAGSGRRHDLRLDPLVLSINPKY
YDALNGVELLKNPDGNLGNPDPLVSPDLIPNRPRIKNSHLLIPTLVAVG
SLVLAFFVYVIMKKKKSKPSNSWCPLPHGTDTNTPKASLPADLCRRFSI
FEKSNATNDEDLIIIGVGFGSVYKQIDGGATLVAVKREITSNOGAKEFELELM
LSKRLHVLVSLGICDEONMVLVYEPHGLTKDHLFRDRKTSDDPLSKSRRTHEIC
IGARGLYLHTGAKTYIITHDKTNNILLDENFVTVKDFGLSRVGPSTASQTHVST
VNGTEGYLDPEYRQVLTKEKSDVYSGVLLVLCRRPIRMOSVPEQADLIRFVK
SVNRGTVDQIIDSLSADITSTSLKFCFIAVRCVQDRGMERPNNVDVVALEFALQ
LHETAKKNDNVSLDMLPSGEGVTTDGEDDLFSRTTGHVGKSTTDDSVLVGDER
SGSSWGVSFSEINPKAR"

6701. 7210

/note="gene_id:MXF12.3"

/evidence=not_experimental

/product="unknown protein"

/codon_start=1

/evidence=not_experimental

/protein_id="BAB10825.1"

/db_xref="GI:10177546"

/translation="MINLEKEAKEQILIRDKRRKIVHHVNTLLVKIYEASAIIDRRH
LFLKPYGPNYFIRKAWLDDTKEYGYVVKDSFKVNPPLKIPNFPAGYLVKELLK
GMSYRDPGTSNRVTVMGGRKIRLSPLMTSYKERPISSGIDLVLGNSDQCQVKKGYLY
AKHLRSAD"

7670. 10111

/note="gene_id:MXF12.4"

/codon_start=1

/evidence=not_experimental

/product="receptor protein kinase-like protein"

/protein_id="BAB10826.1"

/db_xref="GI:10177547"

/translation="MNCNLVFLSLVSVTAGVTAAYHPTDVFLENGDPTSNVNDNSG
RNMTVRSQTLSSNLVNSTSEASYQKAGSIPYMKARIFRSEFTYFPVTPGSGIF
RLRYFYQTVKNSFVNSFFSVKNGFTLNNFSAADLTAKSKQTEFIKEFIIPV
HOTLNLFTPTSKNLLAFVNGIEIVSMDFRYSKGGFNDVLRNVSDDVIDNSTAFE
SVHRLVNGQIVNEVDDSGMFRWLSDDSGNSGSIVNVPVKINTKPTPAYVAPYD
VYATSLMGNSMLNLTGMFLTDVAGTNYLRLHFCETLPQVTRAGQRFISFVIED
KMAKKTVDVIRLSGGPRIPMYLDFSVYVGFESGMIOPELRLDLVPLKDTNQTYDAIL

SGVEILKLNDSGNLARPNPELVSTDDSDPDDSNVTPPIKGRPHVLVILVIVGSGVIG
LATFVIMLLIRQMKKKNKENSIMFKLLKQIYAEIKKNTISFSFSGVKGSGFG
TVYRNLNGRTAVKVKVLDKNGGDDFNEVTSQTSQSHVNIIVSLFCYEGSKRAI
ISEFLEHSGLDQFISRNKSLTPNVTLYGIALGARGLEYLHVGCRTKRVHPDKPON
ILLDDNCPKVADEGLAKCEKRESILSLDTRGTIGYTAPEVVSWMYSGIKSHKSDY
SYGMLVDMIGARKNVTTCNGSTAYFPDWIYKDLDENGDQTWIIIDGINEENKDIYK
KMILVLCIRPCPSDRPPMKVYEMIEGSLDALELPPKPSRHSISTELVLESSLSLDG
QEAKEQTQIDTII"

CDS

10819. 13239

/note="gene_id:MXF12.5"

/codon_start=1

/evidence=not_experimental

/product="receptor protein kinase-like protein"

/protein_id="BAB10827.1"

/db_xref="GI:10177548"

/translation="MICFILLVPSFLVSATAPYKPDVDFLNCCTDVPDPNHGRT
WTOEKNILPKNSDNASFSSVSKESGIPQVPTATIRFSDFTSYSPVSGWRFL
LTYPTYSYSGFDVANSVSTVNDFTLLQNFSAADLTAKASIPESKLKEITVPIY
LTLNLTFRPSNNSLAFVNGIEIVSMDFRYSKGGFDLITNVGSLDDEFIDNSTASET
VHRLNVGGMVDENVDSGMFRWLSDDYEFLLGVSPYMPDVNIYSTKPTPAYVAPY
VYSTCRMGNAODTYLNLNLTWLTVDAGFSLVRLHFCETLSEVNGKQRFVIFLG
NOMAREEMDVIRLSGGPRIPYLDLFRIVYSGESGPRDLRLDHLPLKONPVEYEAII
NGVEILKLNNSGNIILIQNELAPNEPLSNLTPNHVTOQIKGKSSHLLVKIIFAVGP
GTGLATFVYVLMWRQMKRKRKEERVNFKLLNNIYAEIKKNTISFSYIIGKGG
FGTVYGNLSNGRKVAVKVLKDLKGAEDFNEVTSQTSQSHVNIIVSLFCYEGSKR
AIYVEFLENGSLDQFMSRNKSLTQDVTTLTGIALGARGLEYLHVGCRTKRVHFDIKP
ONILLDNLCPKVSDFGLAKCEKRESVLSMDTRGTIGYTAPEVFSRMVGRYSHKSD
VYSFGMLVIDMIGARKEIVETVDSASTYFPDWIYKDLDENGDQTWIIIDGDELTKEEK
EIAKMLIVGLWCIQPCPSDRPPMKVYEMIEGSLDALELPPKPSRHSISTELVLESS
LSDGGBDV"

CDS

join(16412. 16498,16859. 16924,17003. 17038,17085. 17192,
17337. 17461,17541. 17640,17803. 17974,18073. 18191,
18285. 18395,18488. 18632,18712. 18844,18934. 19009,
19232. 19297,19385. 19492,19591. 19698,19873. 20070,
20164. 20373)

/note="gene_id:MXF12.6"

/codon_start=1

/evidence=not_experimental

/product="ABC transporter-like protein"

/protein_id="BAB10828.1"

/db_xref="GI:10177549"

/translation="MGNKLLITGGSKTHGSSYRDPQLQONEDPKANGSBNGLND
LEHGVNEAVNGFRVFCALICQSSIMSOAKPDAGLVIGTIALIGTSLNLLPKF
GMIIDIVSDKVTPEQQTESLIARNAVVIILVIGSICITLAWLFNSASRVV
ARGKMLTFLHMOETADYDVTGGLSRLSDTOIKNAATNLSEALRVNTTALI
GURKDFTSKKLLTALVVPVLSVAVKQFGRVLRSLHTTQAAASAAVAAIESGCA
VFRVRSKESYMYSOYSKKVDETLKLGKQAVLVGLFEGGLNAATLSVTVYVGA
YLITGSMIVGALTSLILSLTVGSSVSSLSYTTAMKAAGSRRVQFLDRVSSMS
SGDGCPCVNGDGVDELNDVWFAYPSRSHMLKISLRLTPGSKVALVPGSGGGKT
IANLIRFYDPLKGIILNGVSLMEISHOYLHQIISIVSQEPILFNCSEENIAYGFD
GEASFTDIENAKMANAHEFIEAFPDKNTVVGGRGLRSGGQKQRIATARALTNP
VLLDDEATSDALDAEVLVDQMDSLMAGRTVLVIAHRLSTVKTADCVASDGEVAE
KGTHDELLSLNGIYTNLVRQLOSSSVTTL"

CDS

join(25348. 26359,26401. 26756)

/note="gene_id:MXF12.7"

/codon_start=1

/evidence=not_experimental

/product="anthocyanin acyltransferase-like protein"

/protein_id="BAB10829.1"

/db_xref="GI:10177550"

/translation="MYNEEMESLKVIVARVTPSNSDSESLTLPUTFDLLWYKLH
AVERIFYKLTADSRFFDSVIVPNKLTSLSSLSHYLPLAGLVMVEPLDKPKIIVYT
PNDAVFTVAESNADFSRLTGKPPTELYPLVPELHVSDSDASAVSFQVTLFPNQ
FCISVNAHVALDQKTTNFKSMARTCKNODSFLPDQLIPVDVTVIKDPMOLDTKI
LNAHVRVAKFTGCKKEPNPKSLKLLSPEIGDFDPVRYTLNLTREDIQRLRERKES
SSGVSAPKRELSTFVIVYALTCLIKARGDPSRPVGVGIVADCDRLSLWVPPVPS
SYFETPMSEGEFLAAARMVSDVEALDENVALKIPILGEGFTTLPSTQVLSVAGSTR
FGYVIGDFGGRPEKVVVWSIDQGEAISFAESRDGSGGVELGFSLKHKHMDVLVDLH
KGLN"

CDS

complement(join(27466. 27879,28046. 29073))

/note="gene_id:MXF12.8"

/codon_start=1

/pseudo

Qy 880 GCATGGTCAAAATCAAGATGACGTCGTATCAGAGGAATCATCGAAGCAACGACGAAATGAGC 939
 Db 28444 GCCAACGCGTATCTTGGACATGCTTGTAAAGCGCTGGAGGA--GATGAGACACAC 28387
 Qy 940 TCGAGTACTTCAGTTTACAGCGGATCCGAGGACTTCTGACGCCGCCGCTCGCGCTA 999
 Db 28386 CGGTCCGTTTCATGCTGCTGATTTTCAGGAACCGGTTAGATCCACCGTCTCTGAGA 28327
 Qy 1000 ACTACTTTGGCAACTGCTT-TCCGTCATCGCTTCCAAAGCAACACATTAAGAGTTAGTT 1058
 Db 28336 TGTATTTCCGGAACGCTGCTGTCGCGATAGGTGCTTTGGTTTACAAGCGCAATGTTT 28267
 Qy 1059 GGGGATAAGGCGTCTTGTTCGACGTTTCAGCTATTTGAGAGAGCCATTGAAAGAGGTTG 1118
 Db 28266 TGGGAGAGATGGGTTGTTTAAATATGGTTCGAATTTCTAAGTGATTCGGTTTGAAGTATTG 28207
 Qy 1119 CACAAAGAAAGGGTTTCTTTCGAGATGCAAAACTTGGTTTATCGGAATCTTAATGGAATC 1178
 Db 28206 GTTTACGAAAGTGGAAACAATTTGCGAATTTGTACATTAATGGAACAAGAGTGTGAAC 28147
 Qy 1179 CTTTCAAAAGATTTCTCGGATACCGGATCGCTAAGTTTCGATTCGTTGCTAGAT 1238
 Db 28146 CAGGTACACAGA-TTGGATCATTGCGCGTTCGAACCAAGTTTGGGCTATACGGGTACAT 28088
 Qy 1239 TTTGATGGGAAAGCCCTGCAAAATTTGACATTACCTCTGTTGATTTATGCAGAATTTGATT 1298
 Db 28087 TTCGGTGGGAAACCATGTTAAGTTCGATTCGCTCCATTGACCGAAACAGGCGGTTT 28028
 Qy 1299 TATGTGATTCAGTCAGGATTTTGAAGAGGTGGAGATGAGTATCATTCCTTAAG 1358
 Db 28027 TCGATGTCGGAGAGAGGAGTCAACCCGAGGTGGAGATCGGCTGTTGTTTGAAGAAA 27968
 Qy 1359 ATTCAATGATGATGATTTGCAAAATCTTTGAAGAGGCTTTTGTCTCTTCGTAGTACT 1418
 Db 27967 TGTGATGGACATATTTATTTATCTGTTTCAAAATGTTTAGAAATGGTCTTTAGGGT 27908
 Qy 1419 CTTTAAT 1425
 Db 27907 CTTTGAT 27901
 RESULT 15
 AB028618 85690 bp DNA PLN 27-DEC-2000
 LOCUS
 DEFINITION Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone:MOD1.
 ACCESSION AB028618 BA00014
 VERSION AB028618.1 GI:5041971
 KEYWORDS
 SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1 clone:MOD1.
 ORGANISM Arabidopsis thaliana
 Eukaryote; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE 1 (sites)
 AUTHORS Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S.
 TITLE Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the 4,251,695 bp regions covered by 90 P1, PAC and BAC clones
 JOURNAL DNA Res. 7 (3), 217-221 (2000)
 MEDLINE 20363099
 REFERENCE 2 (bases 1 to 85690)
 AUTHORS Kaneko,T., Katoh,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-1999) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)
 COMMENT Address for correspondence: kaos@kazusa.or.jp
 For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MOD1
 Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant

protein similarity are described as 'unknown protein'.
 The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Graal-1.3),
 GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
 NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
 SplicePredictor (Volker Brendel, Stanford University, http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi).
 Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/).
 This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions.
 The 5' clone is T13J10 and the 3' clone is T26G12.
 Location/Qualifiers
 1. 85690
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="MOD1"
 /clone_lib="Mitsui P1"
 2123. 8637
 /note="contains similarity to TNP2-like transposon protein
 gene_id:MOD1.1"
 /codon_start=1
 /pseudo
 /evidence=not_experimental
 join(10671..10844,10921..11208)
 /note="contains similarity to unknown protein
 emb|CAA16722.1
 gene_id:MOD1.2"
 /codon_start=1
 /pseudo
 /evidence=not_experimental
 15142..16497
 /note="gene_id:MOD1.3"
 /codon_start=1
 /evidence=not_experimental
 /product="anthocyanin 5-aromatic
 acyltransferase/benzoyltransferase-like protein"
 /protein_id="BAB02518.1"
 /db_xref="GI:11994477"
 /translation="MTLHVETARVPTDYVINSANLHKPLTFFDLPHLLPQVVR
 VFTEPSTDRHFHSILPKLDSLSLIRNPLTGHITWEPNEPKPSLIVSENG
 VLVTIAESDAFSLHSGYQRPUSLHALVPLKPVSDSATAFSIQITLFPNQGSFG
 VAAHVAIDGTSSTFIKAWAQICKQLQSPENLTPSYDRSLIKYPTYLDERKIELV
 RSLKEDQTNIRSLTSLPSSKIGDDVLATLVLSRADLEREQVKNVSPSLHSTFVI
 AYAVATCFVKARGNKDRSVSLLEFGDFDRDLPKLPQTYFGNCMLPVGYNKAAE
 FMEKGPVTAAEIISDLVGLSSRKIETIADTFVEGFSQSWSTQFCTIAGSTPLGVY
 EADFGWGRPVKVDIVSDQGEATAMAEERDESGGVEIGMCLKKTEMDSVSVFFNGLH
 S"
 complement(18639..19994)
 /note="gene_id:MOD1.4"
 /codon_start=1
 /evidence=not_experimental
 /product="anthocyanin 5-aromatic
 acyltransferase/benzoyltransferase-like protein"
 /protein_id="BAB02519.1"
 /db_xref="GI:11994478"
 /translation="MALNVIKSRVSLVNSVEPLVPLTFFDLMLKLNPIERVTVY
 KLTSSRDSFSSITLPLEOSLIVLSHFLPLSGHLAKNPQDPKPHIVIRPKDPTSLT
 VWEADGFSYISSKELRELEPLVPELOVSSDSALLSQITLFPNQGSIGTIVR
 HVMWDKTSKFSKSWAHKGGTTPQDFDLPTVLDRTVINVPAGLSQKIFQLSSYIS
 EEKDYARTLTPPAKEIDNDVVRVLTETVEIDLEKRAKNSTRLDLSLSTFVSY
 AYVLTCKVSKCGDANRFVMTAAADFNRLDPPVPLTYFGNCVLPIDFNGYKATFV
 KDGKYNVGEILSDSVRLGSRNIESIWEVYEDGTGNKLDQNTQVTVTGSNOFCIYS
 GFGWGRPVKTQVMSLYKNNEFSNSARDEIGGLEIGISLKKCNENVLSLFTSDFDIY
 K"
 complement(join(27960..28438,28676..28739))
 /note="gene_id:MOD1.5"

FEATURES
 source

CDS

CDS

CDS

CDS

CDS


```

unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB02520.1"
/db_xref="GI:11994479"
/translation="MRLAGAKSSPEEERRRRSRARANGSVDRPIRFIYAADRNRLLD
PPVLTIFGNCVLPIDFNGYEATIFLGDGFVNGVEILSDSVKGLGSRKSLSESVWEIYE
EGTKMKVGTGKVLTVTGSNQFIYGADFGWGPVNTVMSLYKNNFSMSARRDEIGV
SLKRCMIVFLSLFSGFDN"
29217..31674
/notes="gene_id:MOD1.6"
/codon_start=1
/pseudo
/evidence=not_experimental
/product="Mutator-like transposase"
join(32547..32726,32904..33122)
/notes="gene_id:MOD1.7"
/codon_start=1
/evidence=not_experimental
/product="Mutator-like transposase-like protein"
/protein_id="BAB02521.1"
/db_xref="GI:11994480"
/translation="MDRJCERDPYDDMKVAKRAIEQEMVAMMEGIPKFCPCGGSIV
DTRKDEKRYOCERKFDORTCMHIRKLWDKAMEEVSLSRESVDYNNRKKVLSHEYLI
EEMQKELAKHRAEIVNVSKVLFPNMPAKK"
complement(join(33460..33573,33657..34016,34094..34255,
34330..35070,35486..35721,35799..36093,36146..36256))
/notes="gb|AAC23763.1
gene_id:MOD1.8
strong similarity to unknown protein"
/codon_start=1
/evidence=not_experimental
/db_xref="GI:11994481"
/translation="MPPTRGGCGQRNEIEASAPAKTEKVKAPAEKVEGAPSKAK
VQAPAKKAVQPAKDEKVEPAEQSPQTQTATAMATNAAPTAAPTTAPTAPT
ESPMDDTSFYDALKHIAEIEENNOTDEVEDETETNKELANCAVPVEABERQDGL
TVIEEERSSESDVDNVKSEVDEHEDERDEVDIVKEGSEERTIDEDIANVDME
AMAPOLGMPVSVPKMKLATRCYISEVLKTFADLEHLTHVEKNVFMHPSEKH
IYHLPSTGTHKLMRWMLFLPTASIEKKKEVWVNGVPIRYGIREHALISGFCKAY
PASYOSAGNNFANRYFTGIRREDVTKLMEPEAPSKORLMAALYFLSIIVMP
TKGERASPIDDFCVRSADLTFCKIFPSGRYSFEYMLKSIKLDHENGVVVNTOSP
WVPGFCVLEPLAFEAIPSLREFIEEKGAGCORMCKVFNFRKRIEMKGFLEQIN
HYLGTVIESIIRKAEVPLLAETIGVEDVDKHDVVDSMMKRLGOGREIRFEV
YNEDVQARMAEPNEEVPTAVGPDPDTLVDVMEKLHSLNDKLEALLVIMEIEEQAT
FEAFDEMKAKSQNPDEESTIKENAAAPVVPKVRSTRAKSSNV"
complement(36608..37000)
/notes="gene_id:MOD1.9
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB02523.1"
/db_xref="GI:11994482"
/translation="MDGKTASKFYKAWAHICKHGNIPQDFDLPWVLDRTVINVPAGLE
PKLFQLLPYLSKEKVNARNLMLPAKENINVVRTLELSEANIKKLKEQAKNESTWSD
LLLSFTVVTYKESGSLFCKEAKKSL"
complement(join(40204..40603,40656..40872))
/notes="gb|AAC97247.1
gene_id:MOD1.10
similar to unknown protein"
/codon_start=1
/pseudo
/evidence=not_experimental
complement(join(42135..42706,42772..43162))
/notes="gene_id:MOD1.11
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB02524.1"
/db_xref="GI:11994483"
/translation="MALNMKISRVSHTNSVTPLEPLPTFFDLLWLKHHPIERTVTFY
KLSESSHDSFSTILPLKEQSLALNPLPSGLHIKWDPODKPRPHIIFPDVTSS
LWLSPTLSQDSPAKSLSVRSYVLLSLRFLSLGTLTHHIMDKRTSTNFKHSAWICK

```

```

YCTIPQGFVLTHSDRTVINSVGLEFKILQLLSYLSGVDNDTRTLKSPSAKEIDDD
LVRVTLQLSQENIKKIKERTKNESTRSDLYLSTFVVKAYVLTVCVVKARGGNVDRPIR
FMVAADERNRLDPPVPLTYFGNCVLPIDIFHFCRRMWCEWRGDSK"
44134..44573
/notes="gene_id:MOD1.12
unknown protein"
/codon_start=1
/pseudo
/evidence=not_experimental
complement(46109..47990)
/notes="contains similarity to reverse transcriptase
gene_id:MOD1.13"
/codon_start=1
/evidence=not_experimental
complement(join(48707..49066,49148..49372,49477..49764,
49850..49982,50772..51282))
/notes="gb|AAF06088.1
gene_id:MOD1.14
similar to unknown protein"
/codon_start=1
/pseudo
/evidence=not_experimental
complement(52336..56745)

Query Match          3.6%; Score 61.8; DB 12; Length 85690;
Best Local Similarity 56.7%; Pred. No. 6.1e-05;
Matches 114; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 1199 GATTACCGGATCGCCTAAGTTCGATTGCTATGCTAGATTGTTGGATGGGAAAGCCTGC 1258
      ||| |||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16287 GATAGCGGGTGCACCCGGTTGGGAGTATACGAGGCGGATTCGGGTGGGGAAGACCCGT 16346

QY 1259 AAAATTGACATTACCTCTCTTGCATTATGCAGAAATTGATTATGTCAGTCCAGGGA 1318
      ||| |||| ||| ||| |||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16347 TAAAGTTGATATTGTCTCCATTGACCAAGGAAGAGCGCATCGCAATGGCTGAGAGACGTGA 16406

QY 1319 TTTTGAAGAGGTGTGGAGATTGGAGTATCATTTGCCTTAAGATTCAATGATGATTCATTTC 1378
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 16407 TGAGTCAGGTGGCGTTGAGATTGGAATGCTTTTGAAAAAGACTGAATGATTCGGTTCGT 16466

QY 1379 AAAAAATCTTTTGAAGAAGCGTTT 1399
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 16467 GTCTTTTTCACAATGGTTT 16487

Search completed: November 5, 2001, 16:30:46
Job time: 9616 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:11:26 ; Search time 445.49 seconds
(without alignments)
2262.188 Million cell updates/sec

Title: US-08-894-356C-3
Perfect score: 1605
Sequence: 1 TGTCGACGAATCCATTCA.....TTAAAAAAAAAAAAAAAAAAAA 1605

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues
Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601.*
1: /SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1601.8	99.8	1605	17	AAT37310 Aromatic acyl tran
2	512.4	31.9	1815	21	AAC48422 Arabidopsis thalia
3	493.6	30.8	1813	21	AAC37537 Arabidopsis thalia
4	378.8	23.6	1359	21	AAC48421 Arabidopsis thalia
5	370.8	23.1	1363	21	AAC40007 Arabidopsis thalia
6	361.6	22.5	2412	21	AAC48416 Arabidopsis thalia
7	140	8.7	1364	21	AAC48416 Arabidopsis thalia
8	136.6	8.5	462	21	AAC48415 Arabidopsis thalia
9	132	8.2	1363	21	AAC39553 Arabidopsis thalia
10	106.2	6.6	1520	21	AAC42254 Arabidopsis thalia
11	81	5.0	936	22	AAF58252 Oligonucleotide D1

12	81	5.0	936	22	AAF58254 Oligonucleotide D1
13	81	5.0	936	22	AAF58257 Oligonucleotide D1
14	81	5.0	936	22	AAF58259 Oligonucleotide D2
15	81	5.0	936	22	AAF58262 Oligonucleotide D2
16	81	5.0	936	22	AAF58255 Oligonucleotide D1
c 17	79.2	4.9	936	22	AAF58252 Oligonucleotide D1
c 18	79.2	4.9	936	22	AAF58254 Oligonucleotide D1
c 19	79.2	4.9	936	22	AAF58257 Oligonucleotide D1
c 20	79.2	4.9	936	22	AAF58259 Oligonucleotide D2
c 21	79.2	4.9	936	22	AAF58262 Oligonucleotide D2
c 22	79.2	4.9	936	22	AAF58255 Oligonucleotide D1
c 23	67	4.2	1477	21	AAC39128 Arabidopsis thalia
24	58.6	3.7	1648	21	AAC64785 Chicken leucocytos
25	57.4	3.6	3399	17	AAT05868 Arabidopsis thalia
26	54.8	3.4	1429	21	AAC33965 Arabidopsis thalia
27	53.6	3.3	1498	21	AAC51224 Arabidopsis thalia
28	53.6	3.3	1500	21	AAC36258 Arabidopsis thalia
c 29	53.2	3.3	1427	21	AAC49764 Arabidopsis thalia
c 30	53	3.3	244	22	AAF58238 Oligonucleotide D1
c 31	52.8	3.3	244	22	AAF58238 Oligonucleotide D1
c 32	51.2	3.2	3579	21	AA70099 Plasmodium falcipa
c 33	50	3.1	2427	11	AAQ04107 Human pro-urokinas
c 34	49.4	3.1	2377	13	AAQ20360 Human pro-urokinas
c 35	47	2.9	926	21	AAA61262 Homo sapiens CO122
c 36	46.8	2.9	2214	19	AAV40524 Plasmodium falcipa
c 37	46.4	2.9	1299	21	AAA70111 Homo sapiens CO122
c 38	45.4	2.8	2814	21	AAA88890 Mouse Wnt-3a gene.
c 39	45.4	2.8	2814	21	AAZ47790 Murine Wnt-3a enco
c 40	45	2.8	1043	13	AAQ21651 Macromolecular sys
c 41	44.8	2.8	1474	18	AAT90174 Oil seed rape cyst
c 42	44.6	2.8	729	20	AAK61443 DNA encoding a hum
c 43	44.6	2.8	1672	22	AAF92066 Human PRO1063 cdna
c 44	44.6	2.8	1701	21	AAZ64982 Membrane-bound pro
c 45	44.6	2.8	1701	22	AAF44128 Human PRO4063 (UNQ

ALIGNMENTS

RESULT 1
AAT37310
ID AAT37310 standard; cdna to mRNA; 1605 BP.
AC AAT37310;
XX
DT 06-FEB-1997 (first entry)
XX
DE Aromatic acyl transferase coding sequence.
XX
KW Aromatic acyl transferase; transformatio; anthocyanin pigment;
KW plants; acylation; colour; tone; colouration; colour change;
KW Gentiana triflora; Petunia hybrida; Perilla oclmoldes;
KW Scenecio cruentus; Lavandula angustifolia; ds.
XX
OS Petunia hybrida (Clone pPAT48).
XX
EH Key Location/Qualifiers
FT CDS 67..1413
FT FT /*tag= a
FT FT /product= Aromatic acyl transferase.
XX
XX WO9625500-A1.
PN
XX
PD 22-AUG-1996.
XX
PF 16-FEB-1996; 96WO-JP00348.
XX
PR 30-JAN-1996; 96JP-0046534.
PR 17-FEB-1995; 95JP-0067159.
PR 29-JUN-1995; 95JP-0196915.
XX
PA (SUNR) SUNTORY LTD.
XX

RESULT	2
AAAC48422	
ID	AAAC48422 standard; DNA; 1815 BP.
XX	
XX	
AC	AAAC48422;
XX	
DT	18-OCT-2000 (first entry)
XX	
XX	
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 57422.
XX	
XX	

KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0128845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 21-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147456.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147915.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148604.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150556.

PR 26-AUG-1999; 99US-0150834.

PR 27-AUG-1999; 99US-0151055.

PR 27-AUG-1999; 99US-0151056.

PR 27-AUG-1999; 99US-0151080.

PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.

ID AAC37537 standard; DNA; 1813 BP.
XX AAC37537;
AC
XX
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 17744.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148555.
PR 13-AUG-1999; 99US-0148634.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

QY 1375 TTGGAGAAAGATAAAGAGTT 1394
| | | | |
Db 1435 cttgtgaagagcaagagtt 1454

RESULT 4
AAC48421
ID AAC48421 standard; DNA; 1359 BP.
XX
AC AAC48421;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 57418.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142603.
PR 09-JUL-1999; 99US-0142620.
PR 12-JUL-1999; 99US-0142677.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147250.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148555.

PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PF		PR	09-JUL-1999;	99US-0142820.
XX	25-FEB-2000; 2000EP-0301439.	PR	12-JUL-1999;	99US-0142877.
PR	25-FEB-1999;	PR	13-JUL-1999;	99US-0143424.
PR	05-MAR-1999;	PR	14-JUL-1999;	99US-0143624.
PR	09-MAR-1999;	PR	15-JUL-1999;	99US-0144005.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144085.
PR	25-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	19-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	28-APR-1999;	PR	20-JUL-1999;	99US-0144884.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	04-MAY-1999;	PR	21-JUL-1999;	99US-0145088.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145192.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151338.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0152353.
PR	28-JUN-1999;	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	PR	22-SEP-1999;	99US-0155139.

PR 06-MAY-1999;	99US-0132487.	PR 22-JUL-1999;	99US-0145192.
PR 07-MAY-1999;	99US-0132863.	PR 23-JUL-1999;	99US-0145145.
PR 11-MAY-1999;	99US-0134256.	PR 23-JUL-1999;	99US-0145218.
PR 14-MAY-1999;	99US-0134218.	PR 23-JUL-1999;	99US-0145224.
PR 14-MAY-1999;	99US-0134219.	PR 26-JUL-1999;	99US-0145276.
PR 14-MAY-1999;	99US-0134221.	PR 27-JUL-1999;	99US-0145913.
PR 14-MAY-1999;	99US-0134370.	PR 27-JUL-1999;	99US-0145918.
PR 18-MAY-1999;	99US-0134768.	PR 27-JUL-1999;	99US-0145919.
PR 19-MAY-1999;	99US-0134941.	PR 28-JUL-1999;	99US-0145951.
PR 20-MAY-1999;	99US-0135124.	PR 02-AUG-1999;	99US-0146385.
PR 21-MAY-1999;	99US-0135353.	PR 02-AUG-1999;	99US-0146388.
PR 24-MAY-1999;	99US-0135629.	PR 02-AUG-1999;	99US-0146389.
PR 25-MAY-1999;	99US-0136021.	PR 03-AUG-1999;	99US-0147033.
PR 27-MAY-1999;	99US-0136392.	PR 04-AUG-1999;	99US-0147204.
PR 28-MAY-1999;	99US-0136782.	PR 04-AUG-1999;	99US-0147302.
PR 01-JUN-1999;	99US-0137222.	PR 05-AUG-1999;	99US-0147192.
PR 03-JUN-1999;	99US-0137528.	PR 05-AUG-1999;	99US-0147260.
PR 04-JUN-1999;	99US-0137502.	PR 06-AUG-1999;	99US-0147303.
PR 07-JUN-1999;	99US-0137724.	PR 06-AUG-1999;	99US-0147415.
PR 08-JUN-1999;	99US-0138094.	PR 09-AUG-1999;	99US-0147493.
PR 10-JUN-1999;	99US-0138540.	PR 09-AUG-1999;	99US-0147935.
PR 10-JUN-1999;	99US-0138847.	PR 10-AUG-1999;	99US-0148171.
PR 14-JUN-1999;	99US-0139119.	PR 11-AUG-1999;	99US-0148319.
PR 16-JUN-1999;	99US-0139452.	PR 12-AUG-1999;	99US-0148341.
PR 16-JUN-1999;	99US-0139453.	PR 13-AUG-1999;	99US-0148565.
PR 17-JUN-1999;	99US-0139492.	PR 13-AUG-1999;	99US-0148684.
PR 18-JUN-1999;	99US-0139454.	PR 16-AUG-1999;	99US-0149368.
PR 18-JUN-1999;	99US-0139455.	PR 17-AUG-1999;	99US-0149175.
PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149830.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140353.	PR 31-AUG-1999;	99US-0151438.
PR 23-JUN-1999;	99US-0140354.	PR 01-SEP-1999;	99US-0151910.
PR 24-JUN-1999;	99US-0140695.	PR 07-SEP-1999;	99US-0152363.
PR 28-JUN-1999;	99US-0140823.	PR 10-SEP-1999;	99US-0153070.
PR 29-JUN-1999;	99US-0140991.	PR 13-SEP-1999;	99US-0153758.
PR 30-JUN-1999;	99US-0141287.	PR 15-SEP-1999;	99US-0154018.
PR 01-JUL-1999;	99US-0141842.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0142154.	PR 20-SEP-1999;	99US-0154779.
PR 02-JUL-1999;	99US-0142055.	PR 22-SEP-1999;	99US-0155139.
PR 06-JUL-1999;	99US-0142390.	PR 23-SEP-1999;	99US-0155486.
PR 08-JUL-1999;	99US-0142803.	PR 24-SEP-1999;	99US-0155659.
PR 09-JUL-1999;	99US-0142920.	PR 28-SEP-1999;	99US-0156458.
PR 12-JUL-1999;	99US-0142977.	PR 29-SEP-1999;	99US-0156596.
PR 13-JUL-1999;	99US-0143542.	PR 04-OCT-1999;	99US-0157117.
PR 14-JUL-1999;	99US-0143624.	PR 05-OCT-1999;	99US-0157553.
PR 15-JUL-1999;	99US-0144005.	PR 06-OCT-1999;	99US-0157865.
PR 16-JUL-1999;	99US-0144085.	PR 07-OCT-1999;	99US-0158029.
PR 16-JUL-1999;	99US-0144086.	PR 08-OCT-1999;	99US-0158232.
PR 19-JUL-1999;	99US-0144325.	PR 12-OCT-1999;	99US-0158369.
PR 19-JUL-1999;	99US-0144331.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144332.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144333.	PR 13-OCT-1999;	99US-0159295.
PR 19-JUL-1999;	99US-0144334.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159534.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160757.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160758.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.

[illegible]

Db 434 tgatggacacaccatgagccctcttctctgctccaggtgacagaattggtagatggaatttt 493
Qy 543 CATGGATTACATTTAAACCATGCTGTGCTGATGGTACTTCGACGTGGCACTTTATGAC 602
Db 494 cataggacgtctctatgaacctcgatgagggagcggctactgctgctggaagtctcttac 553
Qy 603 CTCGTGTCGCGAGCTTGTCTGTGGTCCACCTCAATTTCTGTCGCCACCATTCCTTGAACG 662
Db 554 cgcattggtccgaatacttccaaagacagaagagcaaccaaacgaactgactgtgtcttaa 613
Qy 663 AACCAAGGCTCGTACACTCGAGTCAGCTCAAGCTCAACCTCTCTCAACCATCAGATGCACCCGA 722
Db 614 gaattccaccctggttgagcgtctacatccctgaaggatcacggtctctctattcagccttcc 673
Qy 723 ACATGCTAAGTCAGCAACCAACGCTGATGCCGCGCAACCTAGACCCACCTCTTCGCGA 782
Db 674 atatagtcacccg-----atgaattatcagaacctcacgaatctcctcaagg 727
Qy 783 AAGAGTATTCAAGTTCTCCGAGTTAGCAATTTGACAAATGACAAATCAACAGTCAATGCGCAA 842
Db 728 gagaatgtctcttctcatcagaaccataagaatgctttaaacaagggtcaacccaat 787
Qy 843 CTCAGGAGAGAGCCATTTCTCCATTCACATTCCTCAATCACTCTCCGCAACAGTGTGGCTGAGCCGT 902
Db 788 atgcggaacaacctccatctcatctcttccaatcttccaatcgtaaccggttatatgagatgcat 847
Qy 903 CACAGTGTGCGCGCAACCTCAACGCGGAGGACTACACTGTGTACACTGTGTTGCTGATTG 962
Db 848 acaacggcgagggagattacatctctcgcgcgagaaacaagtctgtagagttgctgctgacaa 907
Qy 963 CAGGAAAGAGGTTGATCTCCCAATGCCAGAAAGTTACTTTCGCGCAACCTAAATTCAGGCAAT 1022
Db 908 cagaggaagaatgatccacgccttcacaaggattacttcggaacactgcctctgcttt 967
Qy 1023 TTTACAGTGTGCGCGGAGGTTTGTACTAGCAAGCCGATCGAGTTCGTGCTGTGGGAT 1082
Db 968 gagaacggctgcaaaagctggtgaactattggaatgatcttggattgctgctttgaa 1027
Qy 1083 GATACAACAAGCTCGTGAAGCATGACGCTAAGGCCATTCATGAAGCAACAAAGGAGTG 1142
Db 1028 agtgcatacaagcagtagctgaacatacagcagcgaagaagatctcaaatgattgatcaatg 1087
Qy 1143 GGAGAGCAACCCGAGATCTTTTCA---GTACAAAGATGCTGGAGTGAACCTGTGTGCTGT 1199
Db 1088 gctaaagtctcttatattaccattatgctggtattgtaaccgcatgagcgttatgat 1147
Qy 1200 TGGAGTTTCGCAAGGTTCAAGTTTACGACCTGGATTTTCGATGGGAAAGCCAGAGAG 1259
Db 1148 gggagctcgcgaggtttaaagtagtgggtgaattgggttggaagagagtgac 1207
Qy 1260 TGTGAGGAGTGTTCGAACAATAGTTTGTGATGGAATGCTGATTTGTACCAAGGCAAAA 1319
Db 1208 gcttagaagtggctatgcgcataagttcgaatggaagagatctcttaccgccgggaagaga 1267
Qy 1320 TGGAGGAAGAGCATGATGTGGAGATTAGTTTGGAGCAAAATGCTATGGAGAGGTTTGA 1379
Db 1268 agggaggggaagtagattgagatttggaggtatgctgttccagagtttatggaagctttgga 1327
Qy 1380 GAAAGATAAAGATTCTCT 1397
Db 1328 atcagatgaagagttcat 1345

RESULT 8

AAC48415

ID AAC48415 standard; DNA; 462 BP.

XX

AAC48415;

XX

AC

XX

DT 18-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 57397.

XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

XX

PN

XX

XX

PD

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

XX

PR 25-FEB-1999; 99US-0121825.

PR

PR 05-MAR-1999; 99US-0123180.

PR

PR 09-MAR-1999; 99US-0123548.

PR

PR 23-MAR-1999; 99US-0125788.

PR

PR 25-MAR-1999; 99US-0126264.

PR

PR 29-MAR-1999; 99US-0126785.

PR

PR 01-APR-1999; 99US-0127462.

PR

PR 06-APR-1999; 99US-0128234.

PR

PR 16-APR-1999; 99US-0128714.

PR

PR 19-APR-1999; 99US-0129845.

PR

PR 21-APR-1999; 99US-0130077.

PR

PR 23-APR-1999; 99US-0130449.

PR

PR 23-APR-1999; 99US-0130510.

PR

PR 23-APR-1999; 99US-0130891.

PR

PR 28-APR-1999; 99US-0131449.

PR

PR 30-APR-1999; 99US-0132048.

PR

PR 30-APR-1999; 99US-0132407.

PR

PR 04-MAY-1999; 99US-0132484.

PR

PR 05-MAY-1999; 99US-0132485.

PR

PR 06-MAY-1999; 99US-0132486.

PR

PR 07-MAY-1999; 99US-0132487.

PR

PR 11-MAY-1999; 99US-0132863.

PR

PR 14-MAY-1999; 99US-0134256.

PR

PR 14-MAY-1999; 99US-0134218.

PR

PR 14-MAY-1999; 99US-0134219.

PR

PR 14-MAY-1999; 99US-0134221.

PR

PR 14-MAY-1999; 99US-0134370.

PR

PR 18-MAY-1999; 99US-0134768.

PR

PR 19-MAY-1999; 99US-0134941.

PR

PR 20-MAY-1999; 99US-0135124.

PR

PR 21-MAY-1999; 99US-0135353.

PR

PR 24-MAY-1999; 99US-0135629.

PR

PR 25-MAY-1999; 99US-0136021.

PR

PR 27-MAY-1999; 99US-0136392.

PR

PR 28-MAY-1999; 99US-0136782.

PR

PR 01-JUN-1999; 99US-0137222.

PR

PR 03-JUN-1999; 99US-0137528.

PR

PR 04-JUN-1999; 99US-0137502.

PR

PR 07-JUN-1999; 99US-0137724.

PR

PR 08-JUN-1999; 99US-0138094.

PR

PR 10-JUN-1999; 99US-0138540.

PR

PR 10-JUN-1999; 99US-0138847.

PR

PR 14-JUN-1999; 99US-0139119.

PR

PR 16-JUN-1999; 99US-0139452.

PR

PR 16-JUN-1999; 99US-0139453.

PR

PR 17-JUN-1999; 99US-0139492.

PR

PR 18-JUN-1999; 99US-0139454.

PR

PR 18-JUN-1999; 99US-0139455.

PR

PR 18-JUN-1999; 99US-0139456.

PR

PR 18-JUN-1999; 99US-0139457.

PR

PR 18-JUN-1999; 99US-0139458.

PR

PR 18-JUN-1999; 99US-0139459.

PR

PR 18-JUN-1999; 99US-0139460.

PR

PR 18-JUN-1999; 99US-0139461.

PR

PR 18-JUN-1999; 99US-0139462.

PR

PR 18-JUN-1999; 99US-0139463.

PR

PR 18-JUN-1999; 99US-0139750.

PR

PR 18-JUN-1999; 99US-0139763.

PR

PR 21-JUN-1999; 99US-0139817.

PR

PR 22-JUN-1999; 99US-0139899.

PR

Db 304 atcaacgcgtagagttctcgtcgtcatgctgtagcgttaaccgtcgatgatctaacg 363
Qy 424 GATGAGAGGACCAACCAATTGCAGGACTTGATTCCTTGTATAATAAATCTTGAATTG 483
Db 364 gctgaggacggagcgctgaagtccaaggagttggtctctacaatgggatcttgaaacttg 423
Qy 484 GAAGGGCTTCATCGCCCTCTCTAGCTGTGCAG 516
Db 424 gaaggacttagacagaccactccttgcgtgtccag 456

RESULT 9

ID AAC39553 standard; DNA; 1363 BP.

XX AAC39553;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 25056.

KW Hybridisation assay; genetic mapping; gene expression control;
protein identification; signal transduction pathway;
metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132484.

PR 06-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.

RESULT 10
AAC42254
ID AAC42254 standard; DNA; 1520 BP.
XX
AC AAC42254;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 34861.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-01233180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127452.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

Qy	1040	CAGGTTTGTTACTAGCAAGCCGATCGAGTTCGCTGGTGGGATGATACAAACCGGATCG	1099
Db	1141	cggcagaggttttctcacgtatctaaatggtcgctgatacagcttaaccagagtggtg	1200
Qy	1100	TGAAGCATGACGCTAAGGGCCATTGATGAAGAAACAAAGGAGTGGGAGCAACCCGAAGA	1159
Db	1201	cgctcatcaagacgaggagattcgtagcgttctggtcggaattgggaagcgaatccccaggt	1260
Qy	1160	TCHTTCAGTACAAAGATGCTGGGAGTGAAGTCTGTTGCTGTTGGAGATGTCGCCAAGGTTCA	1219
Db	1261	gtttctctcttggaaatgcgcgacgagcttcggttcactatgggaagctctctccgaggtttc	1320
Qy	1220	AGGTTTACGACGTGGATTTTGGATGGGAAAGCCAGAGAGTGTGAGGAGTGGTTCGAACA	1279
Db	1321	cgatgatataatgatcttcgggttgggaagcgggtggcgggttagaagtggacggctcga	1380
Qy	1280	ATAGGTTTGATGAATGGTGCTATTGTACCAAGGCCAAAAATGGAGGAGAAGCAATTCATG	1339
Db	1381	ataagtttgatgggaagatttcggggttccccggggaagggaagaaacgggactgtgatt	1440
Qy	1340	TGGAGATTAGTTTGGGAAGCAATGCTATGGAGAGGTTGGAGAAAGATAAAGAGTT	1394
Db	1441	tgaagatgggtttctatccggagacatcaagctggatctgaattgaattctcaacgattgaatt	1495

RESULT 11
AAF58252
ID AAF58252 standard; DNA: 936 BP.

DT 24-APR-2001 (first entry)

XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.

XX
PN
WC20010766

PD 01-FEB-2001.
XX

PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 00US-0100350
PR 17-MAR-2000; 00US-0100350

XX XX

XX

XX
XX
XXXXXXXXXXXXX

CC and present an intrinsic electron-conducting property. These materials
CC include each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.

XX
S0 sequence 936 BP: 4 A: 139 C: 10 G: 7 T: 776 other:

Query Match 5.0%; Score 81; DB 22; Length 936;
Best Local Similarity 1.1%; Pred. No. 2.3e-10;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 16:38:25 ; Search time 8904.87 Seconds
(without alignments)
2787.883 Million cell updates/sec

Title: US-08-894-356C-3
Perfect score: 1605
Sequence: 1 TGTCGACGAAATCCATTTC.....TTAAAAAAAAAAAAAAAAAAAA 1605

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : GenEmbl:*
- 1: gb_ba1:*
 - 2: gb_ba2:*
 - 3: gb_ba3:*
 - 4: gb_in1:*
 - 5: gb_in2:*
 - 6: gb_in3:*
 - 7: gb_om:*
 - 8: gb_ov:*
 - 9: gb_pat1:*
 - 10: gb_pat2:*
 - 11: gb_ph:*
 - 12: gb_pl1:*
 - 13: gb_pl2:*
 - 14: gb_pl3:*
 - 15: gb_pl4:*
 - 16: em_ba1:*
 - 17: em_ba2:*
 - 18: em_fun:*
 - 19: em_htgo_hum:*
 - 20: em_htgo_inv:*
 - 21: em_htgo_rnd:*
 - 22: em_htgo_hum1:*
 - 23: em_htgo_hum2:*
 - 24: em_htgo_hum3:*
 - 25: em_htgo_hum4:*
 - 26: em_htgo_hum5:*
 - 27: em_htgo_hum6:*
 - 28: em_htgo_hum7:*
 - 29: em_htgo_hum8:*
 - 30: em_htgo_inv1:*
 - 31: em_htgo_inv2:*
 - 32: em_htgo_other:*
 - 33: em_htgo_rnd:*
 - 34: em_hum1:*
 - 35: em_hum2:*
 - 36: em_hum3:*
 - 37: em_hum4:*
 - 38: em_hum5:*
 - 39: em_hum6:*
 - 40: em_hum7:*
 - 41: em_in:*
 - 42: em_om:*
 - 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_rts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vi:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_vil:*
- 59: gb_vl2:*
- 60: gb_vl2:*
- 61: gb_vl2:*
- 62: gb_vl2:*
- 63: gb_vl2:*
- 64: gb_vl2:*
- 65: gb_vl2:*
- 66: gb_vl2:*
- 67: gb_vl2:*
- 68: gb_vl2:*
- 69: gb_vl2:*
- 70: gb_vl2:*
- 71: gb_vl2:*
- 72: gb_vl2:*
- 73: gb_vl2:*
- 74: gb_vl2:*
- 75: gb_vl2:*
- 76: gb_vl2:*
- 77: gb_vl2:*
- 78: gb_vl2:*
- 79: gb_vl2:*
- 80: gb_vl2:*
- 81: gb_vl2:*
- 82: gb_vl2:*
- 83: gb_vl2:*
- 84: gb_vl2:*
- 85: gb_vl2:*
- 86: gb_vl2:*
- 87: gb_vl2:*
- 88: gb_vl2:*
- 89: gb_vl2:*
- 90: gb_vl2:*
- 91: gb_vl2:*
- 92: gb_vl2:*
- 93: gb_vl2:*
- 94: gb_vl2:*
- 95: gb_vl2:*
- 96: gb_vl2:*
- 97: gb_vl2:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1603.4	99.9	1605	10	E12755	E12755 Petunia hyb
2	1603.4	99.9	1605	12	AB026495	AB026495 Petunia x
c	378.8	23.6	82415	12	AB005244	AB005244 Arabidops
4	135.2	8.4	79976	12	AB008264	AB008264 Arabidops
5	131.8	8.2	118507	13	ATF7J8	AL137189 Arabidops
6	119.2	7.4	83599	12	AB005249	AB005249 Arabidops
7	119.2	7.4	88095	13	ATF13G24	AL133421 Arabidops
c	106.2	6.6	110531	12	AF002109	AF002109 Arabidops

```
c 9 102.8 6.4 88318 13 AP002842 Oryza sat
10 94.4 5.9 2072 14 MTR310831
11 68.8 4.3 111122 13 ATFL11C1
c 12 67 4.2 7218 10 I66494
c 13 67 4.2 74342 12 AB020742
c 14 60.8 3.8 137205 83 AP003279
15 60.8 3.8 151703 83 AP003442
16 58.6 3.7 1648 9 AX025514
17 57.4 3.6 2069 45 E10125
18 57.4 3.6 3399 45 E10126
c 19 55 3.4 148608 14 H0711G06
20 53.2 3.3 1120 6 D1HFERA
c 21 52.8 3.3 152409 84 PFMAL1P1
c 22 51.6 3.2 113880 96 PFMAL3P4
c 23 51.4 3.2 854 54 CNS07DDY
c 24 51.2 3.2 12029 4 AE001373
25 51 3.2 7736 5 AF310889
26 51 3.2 111554 13 AP002460
c 27 50.6 3.2 82360 13 AP000606
28 50.6 3.2 341064 75 AC074226
29 50.4 3.1 2407 5 AF050739
30 50.2 3.1 2263 5 AF251290
31 50 3.1 191034 80 AL356372
32 50 3.1 253305 96 PFMAL3P7
c 33 49.8 3.1 1732 6 DDIECBPRO
34 49.6 3.1 4163 6 DDIBP74
35 49.4 3.1 196490 60 AC005507
36 49.2 3.1 14529 4 AE001397
c 37 49.2 3.1 114226 60 AC004710
c 38 49.2 3.1 196490 60 AC005507
39 48.8 3.0 13859 4 AE001366
c 40 48.8 3.0 83110 84 PFMAL13PB
41 48.6 3.0 397 54 G37798
42 48.6 3.0 2384 96 PFAGAR
43 48.6 3.0 12972 5 AF088979
c 44 48.4 3.0 1124 53 CNS073BM
c 45 48.4 3.0 2265 5 AF070091
```

ALIGNMENTS

```
RESULT 1
E12755 1605 bp DNA PAT 24-JUN-1998
LOCUS Petunia hybrida mRNA for acyltransferase, complete cds.
DEFINITION
ACCESSION E12755
VERSION E12755.1 GI:3251587
KEYWORDS JP 1997070290-A/3.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1605)
AUTHORS Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y.,
Yonekura,K., Mizutani,M. and Kusumi,T.
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 3 18-MAR-1997;
SUNTORY LTD
COMMENT OS Petunia hybrida
PN JP 1997070290-A/3
PD 18-MAR-1997
PR 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO.
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
C12R1:865),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH Location/Qualifiers
FH Key
```


Qy	694	AACCTCTCTCAACCATCAGATGCACCGAACAATCTAAGTCAGCAACCAACGGGTGATGTC	753
Ds	75123	GACCTACTGCTCCCAAGGACCCCAACGAGACCTCCACGCGGAA---GATGCTGCCAAT	75067
Qy	754	CCGGCCAACTAGACCAACCTCTTCGGGAAAGAGATATCAAGTTCTCCGAGTTAGCAATT	813
Ds	75066	CCGACGGTGGAAACCGCGGACCTAGTAGAAGAATCTTCAGTTCTCGGACTTTCGCTGTC	75007
Qy	814	GACAAATCAAGTCAACAGTCAATGCCAATCTCAGAGAG-----ACGCCAATTCTCC	864
Ds	75006	CACAGTCAAGTCAAGCTAAATTCAGTAAATCCATCGGACAGTTTCAAAACCATCTCA	74947
Qy	865	ACATTCCAATCACTCTCCGACAGCTGTGGCTAGCCGTACACAGTGGCGGCAACTCAAG	924
Ds	74946	ACATTCCAATCCCTGACATCACATCATCTGGCGCCACGTCACCTTAGCGGTGGACTCAA	74887
Qy	925	CCGAGGACTACACTGTGACACTGTGTTGCTGATTCAGGAGAAAGGTTGATCCTCCA	984
Ds	74886	CCGAGGAGCATAACTATTTTCCACGCTCTCCGCGACTTCGCCGCCGCTTGATCCTCCA	74827
Qy	985	ATGCCAGAAAGTTACTTCGGCAACCTAATTCAGGCAATTTTTCAGAGTCACGCGCGCAGGT	1044
Ds	74826	ATGCCGAGGAATATTTCCGAAACTTGATTCAGCGATCTTTCACCGGAACAGCGCGGGG	74767
Qy	1045	TTGTTACTAGCAAGCCCGATCGAGTTCGCTGGTGGATGATACAAACGATCGTGAAG	1104
Ds	74766	CTACTAGCGCGCACCGACCGAGTTCGGAGCTTCGGTGATCCAGAAACATCGCGCA	74707
Qy	1105	CATGACCGTAAAGCCATGTATGAAGAAACAAAGAGTGGGAGACACCCGAAGATCTTT	1164
Ds	74706	CACGACCGAGTGTGTCGACGCGCACGTAAACAGCAGTGGGAGAAATGCCCGAAGATTTT	74647
Qy	1165	CAGTACAAAGATCTGGAGTCAACTGTGCTGCTGTCGAGTTCGCAAGTTCGAAGTT	1224
Ds	74646	CAATTCAAAGACCGGAGTGAATGTGTCGGTGGGAAGCTCACCGAGTTTCGTTGTT	74587
Qy	1225	TACGACTGTGATTTTGGATGGGAAAGCAGAGAGTGTGAGAGTGTGTCGAACAATAGG	1284
Ds	74586	TACGAGGTGGATTTTCGGGTTTGGAAACCGGAACCGGTAGGACGGTCAATTAACCG	74527
Qy	1285	TTTGATGGAATGTTGATTTGATACAGGCAAAATGAGGAAGAACGATGATGAGG	1344
Ds	74526	TTTAATGGGATGATGATTTGATACCAAGGTAAAGCAGGACATTAAGTATTGATGAGG	74467
Qy	1345	ATTAAGTTTGAAGCAAAATGCTATGAGAGGTTGGAGAAAGATAAGAGTT	1394
Ds	74466	ATACTCTGAAGCTTCTGTTATGAGAGAGCTTGTGAAGCAAGAGTT	74417
RESULT	4		
LOCUS	AB008264		
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone:MBD2.		
ACCESSION	AB008264	BA000015	
VERSION	AB008264.1	GI:2618599	
KEYWORDS	Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui p1 clone:MBD2.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
AUTHORS	1 (sites)		
TITLE	Nakamura,Y., Sato,S., Kaneko,T., Kotani,H., Asamizu,E., Miyajima,N. and Tabata,S.		
JOURNAL	Structural analysis of Arabidopsis thaliana chromosome 5. III. Sequence features of the regions of 1,191,918 bp covered by seventeen physically assigned p1 clones		
MEDLINE	DNA Res. 4 (6), 401-414 (1997)		
REFERENCE	98162728		
AUTHORS	2 (bases 1 to 79976)		
	Nakamura,Y.		

TITLE	Direct Submission
JOURNAL	Submitted (24-OCT-1997) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakam@kazusa.or.jp, tel:81-438-52-3935, fax:81-438-52-3934)
COMMENT	Address for correspondence: kaos@kazusa.or.jp For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/seqd/graph.cgi?c=MBD2 Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Graal (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Graal-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://greml.in.zoel.iastate.edu/cgi-bin/sp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is MJB21 and the 3' clone is MRD20. Location/Qualifiers 1..79976 /organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /chromosome="5" /clone="MBD2" /clone_lib="Mitsui p1" complement(1..218) /note="CDS is reported in Acc# AB007647 gene_id:MJB21.20" /number=1 /evidence=not_experimental /product="U2 snRNP auxiliary factor, small subunit" join(5306..5761,5932..6828) /note="gene_id:MBD2.2" /codon_start=1 /evidence=not_experimental /product="N-hydroxycinnamoyl/benzoyltransferase-like protein" /protein_id="BAB09184.1" /db_xref="GI:9758571" /translation="MDSSSVKIVSKSVKPKTLPEESKOPYLSPWDYAMLVSQYIQ KGLLFHKPLDSDITLLEKLDLSAVTLVHFYLAGRLSLATTEKPKRSVFDKNDSPGAGFIYATSDLCIKDIVGAKYVPSIVQSFDDHKA VNHGHTMSLSVQVTELVDCI FIGLSNHAAGDGTAEWKFPTAWSEIFQGESQNDLCLKNPVLKRYIPEGYGLF SLPSHPDEPIRTYSPILKRMFCFSETIRMLKTRVNOICCTTSSPQSITAVTW RCITRARRLPDRETSCTRNAADNRMYPLHKDYFGNCLSLARTAKAGELLENDLG FAALKVHQAHEHTSEKVSQMDWLKSPYIHDLRFPMVSVMGSSSPFNKYGCEFG LGKGVTLSRGYAHKFDGVSAYPGREGGSDIDVCLVPEFMEALDESDEERFLSVSL" 8080..10095 /note="gene_id:MBD2.3" /codon_start=1 /evidence=not_experimental /product="CHP-rich zinc finger protein-like" /protein_id="BAB09185.1" /db_xref="GI:9758572" /translation="MSLYFEGHEHHVSIKHRDGLSCDADRSGFDVISCGECKFTVHKRCVFMFDQIEFDHPSHGHLKLLTTPADDDTQDCCHLCGRKTKRLLYHSDCKLN VDDICIDYICARSPKLPMWHDPLIKVDLGANNPCDFCNESGIDPCRCRPMIHER CASVDFSPETTHPSHVHPLKLNPNGLPDTNLNCHICGDTAGNLLYHCDICRFLNMR CAIREPTVALSNMKVHEHTLMPRLISFVCDACGMKDRAPYVCQCDYFIHOE CAKLPRIHVNHHDHRSYKYPGLGCGEWCVEEDMSYGAYSCSLCQYAMHSIC ATRKDWGDELGVLEEDIEDIEPFKRNDDNTITHTHEHNLKSDGITLAKSILCVAC VCPGIDTFINSESSCFVLKGRHFLSPVLVLCQNRQNTCTNACQ QVECKGFIYSSFPKTVYREKFDYDQSSITVPEFHSHDHLHLFLKLRGNVKTCCCK GIVQKEYAIGCTKCNVFLDFRCATLPLTVRLPRYDDHPLTLCYGDEKASGKCWDCICE

[illegible]

```

/number=1
8214..8420
/gene="F7J8_10"
/number=2
complement(8811..8995)
/gene="F7J8_20"
/number=1
complement(join(8811..8995,9120..10082,10181..10291,
10608..10852,10996..11147,11461..11559))
/gene="F7J8_20"
8811..11559
/gene="F7J8_20"
complement(join(8811..8995,9120..10082,10181..10291,
10608..10852,10996..11147,11461..11559))
/gene="F7J8_20"
/notes="strong similarity to lac110 laccase, Populus
trichocarpa, EMBL:PTV13773"
/codon_start=1
/product="laccase-like protein"
/protein_id="CAB69832.1"
/db_xref="GI:6759427"
/translation="MPRLHHVLSNQAFLLVLLFSSIASAAVVEHVLHIQDVVVKPLCK
EQIIPAAAGSLPGPTINVREGDTLVVNVINNSTYNVTIHHGVFOLKSVWMDGANMIT
QCPQPGNYFTQDITGQEGTLLWHAHVNLRAITLHGALVIRPSGRYPFPKPYKE
VPIVFOQWMDVRLQLRPAVSDAYLINGLAGDSYPCSENRMFNKLVQGTLYLLR
IVNAALNTHLEFKIANHNVTVAVDVYSTPLTDVMTLTPGQTVDAALLTADQAIKY
YMATLPIYSAIGIPTPDKPTRGILVYOGATSSSPAEPLMPVNDNSTAHRFTSNIT
SLVGGPHWTPVRHVDKMFITMGLGLDPCPAGTKICIGLQRTAGSLNNRTFMIPER
ISMQEAYNTSGITDDFPNPKFDYTFEQRTNNDKMKMFPERKTSVKKIRFNS
TVEIVLQNTAISPSHPMLHGFNFYVLGYGFGNYDPIRDARKLNLFNPMQHNVTGV
PPGGWVLFRTANNPGVWLFCHMDAHLPIGIMSAFIVQNGPTPETSLSPPSNLPQC
TRDPTIYDSRTNTDLSV"
complement(8996..9119)
/number=1
complement(9120..10082)
/gene="F7J8_20"
/number=2
complement(10083..10180)
/number=2
complement(10181..10291)
/gene="F7J8_20"
/number=3
complement(10292..10607)
/number=3
complement(10608..10852)
/gene="F7J8_20"
/number=4
complement(10853..10995)
/number=4
complement(10996..11147)
/gene="F7J8_20"
/number=5
complement(11148..11460)
/number=5
complement(11461..11559)
/gene="F7J8_20"
/number=6
complement(join(13626..13810,13898..14866,14987..15097,
15387..15631,15775..15926,16131..16229))
/gene="F7J8_30"
13626..16229
/gene="F7J8_30"
complement(join(13626..13810,13898..14866,14987..15097,
15387..15631,15775..15926,16131..16229))
/notes="strong similarity to lac110 laccase, Populus
trichocarpa, EMBL:PTV13773
Contains Short-chain dehydrogenases/reductases family
signature AA184-212"
/codon_start=1
/product="laccase-like protein"
/protein_id="CAB69833.1"
/db_xref="GI:6759428"

```

```

/translation="MPRVHHSNQAFLLVLLFSSIASAAVVEHVLHVKDVVVTPLCK
EOMIPVNGSLPGPTINVREGDTLVVNVINNSTYNVTIHHGVFOLKSVWMDGANMIT
QCPQPSNNITQDITGQEGTLLWHAHVNLRAITLHGALVIRPSGRYPFPKPYKE
VPLIFQWMDVRLQLRPAVSDAYLINGLAGDSYPCSKRMFNKLVQGTLYLLR
IINAALNTHLEFKIANHNVTVAVDVYSTPLTDVMTLTPGQTVDAALLTADQAIKY
YMATLPIYSAIGIPTPDKPTRGILVYOGATSSSPAEPLMPVNDNSTAHRFTSNIT
SLVGGPHWTPVRHVDKMFITMGLGLDPCPAGTKICIGLQRTAGSLNNRTFMIPER
ISMQEAYNTSGITDDFPNPKFDYTFEQRTNNDKMKMFPERKTSVKKIRFNS
TVEIVLQNTAISPSHPMLHGFNFYVLGYGFGNYDPIRDARKLNLFNPMQHNVTGV
PPGGWVLFRTANNPGVWLFCHMDAHLPIGIMSAFIVQNGPTPETSLSPPSNLP
QCTRDPTIYDSRTNTDMSY"
complement(13626..13810)
/gene="F7J8_30"
exon
Query Match 8.2%; Score 131.8; DB 13; Length 118507;
Best Local Similarity 52.9%; Pred. No. 1.4e-20;
Matches 283; Conservative 0; Mismatches 252; Indels 0; Gaps 0;
QY 860 TCTCCACATTCACATCTCTCCGACACGCTGTGGCTAGCCGTACACGTCGCGCCCAAC 919
DB 80827 TTTTCATCGTTCCAACTCTCTCAGCGCTCAGCTATGCGGATCCGTTACACGAGCGGAATC 80886
QY 920 TCAAGCCGAGACTACACTGTGTACACTGTGTTCGTGCTGATTGCAGGAAAAGGTTGATC 979
DB 80887 TCGATCCGAGCAAGACGACGAGCTTTTCGAATGCGCGTTAATTGCCGCGCACCGCTTGAGC 80946
QY 980 CTCCAATGCCAGAAAGTTACTTCCGCAACCTAATTCAGCAATTTTCACAGTGACCGCGG 1039
DB 80947 CGAAGATGGATCCGTACTACTTCCGGAACCGCATACAGAGCATACCGGCTTGGCGCTCG 81006
QY 1040 CAGGTTTGTACTAGCAACCGCGATCTCGCTGTGGTGGGTGATGATACAAAGCGATCG 1099
DB 81007 CGGAGATCTGCTAAGCAAGATCTCAGGTGTCGCCCAACAGTTTACACGAGAACGTGG 81066
QY 1100 TGAAGCATACCGCTAAGGCCATTTGATGAAGAAACAAAGAGTGGGAGAGCAACCCGAGA 1159
DB 81067 TGGCGCAGCAGCAGCGACGCTCCCGCGTGGAAATCGCCGCTTGGAAAGCGATCCGAGC 81126
QY 1160 TCTTTTCAGTACAAAGATGCTGGAGTGAACCTGCTGCTGCTGGAAGTTCGCCAGTTTCA 1219
DB 81127 TGTTCCTCTCCGAATTCAGATGGAGCTTCAGATCAGATGGGAGCTCGCGAGATTCC 81186
QY 1220 AGGTTTACGACGTGGATTTTGTGGGAAAGCCAGAGAGTGTGAGGAGTGTTCGAACA 1279
DB 81187 CAATGTACACAATGATTTTCGGATGGGAAACCCGTTAGCTGTGAGAACGCGCGGAGCA 81246
QY 1280 ATAGTTTGTGATGGTGTGATTTGTACAGGCAAAAATGGAGAAAGCAAGCATTTGATG 1339
DB 81247 ATAAATTCAGCGGAGATCTCGCGCTTTCCCGGTAGAGAAGAAACGGAAGCGTGGATC 81306
QY 1340 TGGAGATTAGTTTGGAAAGCAATGCTATGAGAGAGTTGGAGAAAGATAAAGAGTT 1394
DB 81307 TGAAGTAGTTCTGGCGCCGAGACTATGCTGGGATTTGAGAAGCATGCTGAGTT 81361
RESULT 6
AB005249 83599 bp DNA PLN 27-DEC-2000
LOCUS Arabidopsis thaliana genomic DNA, chromosome 5, PL clone:MXM12.
ACCESSION AB005249 BA000015
VERSION AB005249.1 GI:2264321
KEYWORDS
SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
clone:MXM12.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (sites)
AUTHORS Sato,S., Kotani,H., Nakamura,Y., Kaneko,T., Asamizu,E., Fukami,M.,
Miyaajima,N. and Tabata,S.
TITLE Structural analysis of Arabidopsis thaliana chromosome 5. I.
Sequence features of the 1.6 Mb regions covered by twenty

```


28500..28883)
/note="contains similarity to heparanase
gene_id:MM12.7"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB09947.1"
/db_xref="GI:10176717"

/translation="MGFRVCVIVVFLGCLLVPEKTMQEMKRASIVIOGARRVCETD
ENFYVDLDMWPHDKNYDCQPMGYSVINDLTPLLTKAKFKPKPIRIRIGSLQD
QVIVDGNLTKPCRPPOKNSGLFGSKGLHMRKDELNSFLATGAVTFEGLNALR
GRHLRGRAMGDHDIHQDLFTVTVSGIVDSWFEFNGNELSGVGSVAVSYLYGK
DLVLVDKINKYNSWHLKPIKLVAPGGFYEQWYTKLLEISGSPVSDVTHHYNELG
SGNDPALVKKIMDSYLSQVSKTFKDVNQTQEHCPWASPMWGESGAYNSGGRHVS
TFIDSFWMLDGLNSARHNTKVCYRQTLVGGFYGLLEKGTVPNPDIYSALLMHLRMG
KGLAVQDGLPOLRVYAHCSKGRAGVTLLINLSNQDFTVSVSNGINVLNAESRK
KSLDLTLKRPSTMGSKADGYLNRREYHLTPENGVLRSKTMVNLNGLSLKPTATGDI
PSLEPVRYSVNSPLNVLPLSMFSFVLPNFDASCS"
complement(29080..29607)
/note="gene_id:MM12.8
unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB09948.1"
/db_xref="GI:10176718"

/translation="MLQBPVALSFRPNSFRRSRSMETGVDRDRGWTQHLKAREGDL
KAVKLLDQGDVADWALAGPKSKGMTPHLAAKGHIEVMDDLLEGRANMEARTSGAC
GWTPLHAAAKRKRKRAVFLVGNAGFLPDDITDSRFNPVQVCHGLEWAYEERKKLSE
DTSLSGDTSCSSAN"
Join(31233..31706,31923..32819)
/note="gene_id:MM12.9"
/codon_start=1
/evidence=not_experimental
/product="anthranilate
N-hydroxycinnamoyl/benzoyltransferase-like protein"
/protein_id="BAB09949.1"
/db_xref="GI:10176719"

/translation="MDSISSEVKIVSQCFVKPKTIPEKWKPEYHLSPLDLVMSHHYL
ONGLLFLKSDDATKTKDFMETWMLQKLRDSLAETLVHFYPLAGRLSLTKLTONPRSVSE

Query Match 7.4%; Score 119.2; DB 12; Length 83599;
Best Local Similarity 49.2%; Pred. No. 1.3e-17;
Matches 476; Conservative 0; Mismatches 463; Indels 29; Gaps 5;

QY 507 AGCTGTGCAGCTCACCAAGCTCAAGACGGCTCACCATGGGATTAGCATTTAAACCATGC 566
DB 31913 ATCTTTTACAGTAACAGAAATGGTAGATGGAGTTCATTTGATTTCAATGATCAACGC 31972
QY 567 TGTGCTGGATGTTACCTGACCTGGACATTTATGACCTCGTGGTCGGAGCTTTGCTGTG- 625
DB 31973 TGTGGGAGATGAGGTTCTTGTGGCATTTTTCACACTCTTTGTCGGAGATTTTCAATGC 32032
QY 626 ---GGTCCACCTCAATTTCTGTCGCCACCATCTCTTGAACGCAACCAAGGCTGTAACACTC 682
DB 32033 ACAAGAAACCGATAACTCTCTTCAAGAATCTCTCCAGTCTTTCACGTTGGTTCTTAA 32092
QY 683 GAGTCAAGCTCAACCTCTCTCAACCATCAGATGCACCCCAACATGCTAAGTCAGCAACCA 742
DB 32093 AGGTACGCT---CCTGTGTACAACTTCTTCACTCACTCTGATGAATTCATTAGCCG 32149
QY 743 ACGGTGATCTCCGGCAACAGTAGACCCACCTCTTCCGGAAGAGATTTCAAGTTCTCCG 802
DB 32150 ATTGGAATCACCGTCT-----TTGAAGGAGAGAAATATTCATTTCTCGT 32193
QY 803 AGTTAGCAATTCACAAATCAAGTCAACAGTCAATGCCAACTCAGAGAGACCCCAATCT 862
DB 32194 CTGAACAATAACATCATTTGAATCAAGGCTAATGAAGAAAGCAGCAACCAACAAATTT 32253
QY 863 CCACATTCCAATCACTCTCCGACACAGTGTGGCTAGCCCTCAGCTCAGCTCGCCCACTCA 922
DB 32254 CTCTTTTCAAGGTTAGCCGATTTATGTGGAGATGCATTAAGAGCAAGAAACTTAC 32313
QY 923 AGCCCGAGGACTACACTGTGTACACTGTCTTGTGCTGATTGCAGGAAAGGGTTGATCCTC 982

DB 32314 CTTATGATCATGAAATTCGTTGTAGCCTTGGCGCTAACAAATGGACAAATTTGGATCCAC 32373
QY 983 CAATGCCAGAAAGTTACTTTCGGCAACCTAATTCAGGCAATTTTTCAGAGTGACCG---CGG 1039
DB 32374 CATTTGCTTTTGGAGCTATCTTGGTAATTTGCTTCTCTCTAAAGTCCAAAGACCGTTACAT 32433
QY 1040 CAGGTTGTTACTTAGCAAGTCCGATCGAGTTCGCTGGTGGGATGATACAAACAGCGCATCG 1099
DB 32434 CCGGTGAGCTACTAGAAAATGATTTGGGATGGCTCTTTGAAAATGCATGAAGCTGTGA 32493
QY 1100 TGAAGCATGACGCTAAGGCTATGATGAAGAAACAAAGAGTGGGAGAGCAACCCGAAGA 1159
DB 32494 TTGGAACACTAGTGAAGTGGTCTCGGAGACCAATCAAAAACTGGTTGAAATCATCTTATG 32553
QY 1160 TCTTTTCAGT---ACAAAGATGCTGGAGTGAACCTGTCTGCTGTTGGAAGTTCGCCAAGGT 1216
DB 32554 TTTTCCATTTAGAAAACTTTTGGGAGCGATGGTGGTTCCATTTGGAAGCTCTCCGAGGT 32613
QY 1217 TCAAGGTTTACGACGCTGGAATTTTGGATGGGAAAGCAGAGAGTGTGAGAGTGGTTCGA 1276
DB 32614 TTAAGATGATGAATCGGAGTTTGGGATGGGAAAGCGGTTGCGGTTAGAAGCGTTATG 32673
QY 1277 ACAATAGGTTTATGATGAATGCTGATTTGTACCAAGCGCAAAATGAGGAGAAAGCATTTG 1336
DB 32674 GCGGTAAGTTTATGATGGGAAAGATATAGCATATGCAAGAAAGAGAGGAGGAGAACCATAG 32733
QY 1337 ATGTGAGAGATAGTTTGGAAAGCAAAATGCTATGAGAGAGTTTGGAGAAAGATAAAGAGTTC 1396
DB 32734 ATTGGAAGTATGCTCTCTTCCAGAGTTTATGGAAGCTTTGGAATCAGATCAAGAAATCA 32793
QY 1397 TCATGGAAGTCTGCTTAATTTGCTTAGCTTGGACTCAACTGGCTTACACTTTTATTATGAGC 1456
DB 32794 TGTCCGTTGCT 32853
QY 1457 TGTATGA 1464
DB 32854 TGTCTGTTA 32861

RESULT 7

ATF13G24 88095 bp DNA PLN 10-DEC-1999
LOCUS Arabidopsis thaliana DNA chromosome 5, BAC clone F13G24 (BSSA project).
ACCESSION AL133421
VERSION AL133421.1 GI:6562294
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 88095)
AUTHORS Bevan, M., Van Der Schueren, J., Chuang, Y.J., Voet, M., Robben, J., Volckaert, G., Bancroft, I., Mewes, H.W., Lemcke, K. and Mayer, K.F.X. Unpublished
JOURNAL 2 (bases 1 to 88095)
REFERENCE EU Arabidopsis sequencing, project.
AUTHORS Direct Submission
TITLE EU Arabidopsis sequencing, project.
JOURNAL Biochimie, Am Klopferstritz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mpg.de, mayer@mpg.de, mayer@mpg.de, Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk
COMMENT Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/tha1/.
FEATURES Location/Qualifiers
source 1..88095
organism "Arabidopsis thaliana"
variety "Columbia"
db_xref="taxon:3702"

gene	/chromosome="5" 445..5396 /gene="F13G24.10" complement(445..585) /number=1 complement(join(<445..585, 585,727..888,1047..1145,1228..1293, 1410..1520,1618..1719,1847..2046,2155..2220,2394..2565, 2966..3025,3257..3347,3554..3627,3851..4026,4311..4443, 4524..4883,5052..5396)) /note="similarity to GYMNOS/PICKLE mRNA, Arabidopsis thaliana, EMBL:AF185578" /codon_start=1 /product="putative protein (fragment)" /protein_id="CAB62593.1" /db_xref="GI:6562295" /translation="MEYSEQRDRDEATRLAFAEKRKQSFDSSENPOQODFLAKCR KLDSNDVCQEGYRVEPVRLSTSPYEFKRVLEICSPDPSVTVQLQGRCPPE QECRLQRKLSDAIPHYQNDGGAGYKKINDYKVSGLRKSASVEVGPWKT LAVVEKQSYISQKQCPLEPYHTEBKVEQLIETLPRLKLVNALLPQLDGLRFLRR GGRCFIADENGLGTLQAIAGHYFISEGILVYCPALRFTWAEELERWLPCLPSD VHLVFGHQNDPAYLPWRPKVYVTSYKMLQHLRTMLEREWALLIVDESHLRGSKKS DPPEIKTVLDVAEKVKHILLSGTPSVSRPGLLGKDKVEPAKTYCEVLVVGQGTIF QDSKGTPLLEILLNOTVMIRLKHLLTQLPPKRRQIVTILLKSDIALAMAVIS EAKQKQGAIEVTEKSHEDPQNRNGSNEAGHVNNAESDGNPSKQNLGKLSYQQL GIKLSAFARWLSLPLSLGLDYTPEDIDGRSRTKRVVFAHHHKVLDGTOEFCDKG IGFVLDGTTLPDRQLAVOSFQSPSEVKIATIGVEAGVGGLDFSAQNTVVFLELPT PSLLIQADEADRAHRGQTSVANNVTFCAKDTMDESNNQNLKHLRISSTVDGYDKGT EIEGTGASIFKPAESERREVELEGQSESNVVAADKIVESCDDPGTGTDTVSNTIDLDK DMTSQLELLEVCPEVSGSGSRSSGTISITMLAQENQHHPKN" complement(586..726) /number=1 complement(727..888) /number=2 complement(889..1046) /number=2 complement(1047..1145) /number=3 complement(1146..1227) /number=3 complement(1228..1293) /number=4 complement(1294..1409) /number=4 complement(1410..1520) /number=5 complement(1521..1617) /number=5 complement(1618..1719) /number=6 complement(1720..1846) /number=6 complement(1847..2046) /number=7 complement(2047..2154) /number=7 complement(2155..2220) /number=8 complement(2221..2393) /number=8 complement(2394..2565) /number=9 complement(2566..2965) /number=9 complement(2966..3025) /number=10 complement(3026..3256) /number=10 complement(3257..3347) /number=11 complement(3348..3553) /number=11 complement(3554..3627) /number=12	intron	complement(3628..3850) /number=12 complement(3851..4026) /number=13 complement(4027..4310) /number=13 complement(4311..4443) /number=14 complement(4444..4523) /number=14 complement(4524..4883) /number=15 complement(4884..5051) /number=15 complement(5052..5396) /number=16 5773..7458 /gene="F13G24.20" complement(5773..7458) /note="similarity to mature parasite-infected erythrocyte surface antigen, Plasmodium falciparum, EMBL:AF056936" /codon_start=1 /product="putative protein" /protein_id="CAB62594.1" /db_xref="GI:6562296" /translation="MSEENISQVDDKCLDDANPMEISVETIPVVDVEVQDDDLVHP AETNHAALKRYSTRVLDDKTKGAQVOTRYRQNTSTHDLCKHGKREDDLVTKPKP VAKKNVDSGLCKGETLRKSLGNVSKPKDKSSLOAAKREAAGEVVKSCDGLHVKKSET KTSFSPSPVVRVVKTNLVNKRASRISQNKSPKEDLSKNLKNKEKGKIVPEVRCDYV LEKTDLEVRKVRISSENKSKEDTLNKKERAKIDEPVCDVLEKTSLDLDAQKVSRISE NKNSKEERLNKLNKENTIDEVDPDAVEKTLVYVESSEVKKKMTSKVKISST QQSSERKIRKPKGKSLSLPPLPPSEVVTGSDPRPIRQTSRSTSLPEKKQSGA NLVTNPKPESKIRPKIGRLGTPPPPTKQMMFKGKVLPEKPEDSTTTSIKFKKIV VQPEKLRSTDVNNKKSLSKDKREGVGKINGEKREKVLVLRHKVEVKKKLQTLFNVI EETVNLKEVRKSKYALVGAPETVISLDNNRTSQKKKTQSKSTSSQVVEGC" complement(5773..7458) /number=1 11145..11303 /gene="F13G24.30" /number=1 11145..13478 /gene="F13G24.30" join(11145..11303,11400..11436,11532..11715,11814..11962, 12037..12273,12372..12561,12665..12790,12901..13000, 13095..13478) /gene="F13G24.30" /note="similarity to heparanase precursor, Homo sapiens, EMBL:AF155510; Contains Serine proteases, trypsin family, active sites AA293-304" /codon_start=1 /product="putative protein" /protein_id="CAB62595.1" /db_xref="GI:6562297" /translation="MAQEMKRASIVIOGARRVCETDENFVCATLDWHPDKCNVDQCP WGYSSVINMDLTRPLTKAKFKAPLRIGSGLOQVYDVGNLTKPCRPQKMNNG LFGFSGCLHMRWDELNSFLTATGAVVTFGLNALGRHLRGLKAWGAWDHIINTQDF LNTVSKGYVIDSWFNGNELSGSGVSGASVSAELYGKDLIVLKDVKNKVNSWLHKPI LVAPGGFYEQOWTKLLEISGPSVDVVTTHIYNLGSNDPALVKKTMDPSYLSQVSK TFKDVGOTIQEHGFWASFPWYSGGAYNSGGRHVSDTFIDSFYLDGLGSHARNTKV YCRQTLVGGFYGLLEKGTFFPNPDYISALLWRLMKGKVLAVQTDGDPQLRVYAKCSK GRAGVTLLILNLSQSDFTVSVNSGNNVLNAESRKKSLDLTLKRFPFISWIGKASDG YLNREEYHLTPENGVLRSKTMVNLNGKSLKPTATGDISLEPVLRSVNSPLNVLPSMS FIVLPNFDASACS" 11304..11399 /gene="F13G24.30" /number=1 11400..11436 /gene="F13G24.30" /number=2 11437..11531 /gene="F13G24.30" /number=2 11532..11715
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron	
exon		exon	
intron		intron</	

	/gene="F13G24.30"	
	/number=3	
intron	11716..11813	
	/gene="F13G24.30"	
	/number=3	
exon	11814..11962	
	/gene="F13G24.30"	
	/number=4	
intron	11963..12036	
	/gene="F13G24.30"	
	/number=4	
exon	12037..12273	
	/gene="F13G24.30"	
	/number=5	
intron	12274..12371	
	/gene="F13G24.30"	
	/number=5	
exon	12372..12561	
	/gene="F13G24.30"	
	/number=6	
intron	12562..12664	
	/gene="F13G24.30"	
Query Match		
Best Local Similarity 7.4%; Score 119.2; DB 13; Length 88095;		
Matches 476; Conservative 0; Mismatches 463; Indels 29; Gaps		
Qy	507 AGCTGTGCAGTCCACCAAGCTCAAGGACGGCGTCACCATGGGATTAGCATTTAACCATGCC	566
Db	16508 ATCTTTTCAGTAACAAGAATTTGGTAGTGGAAGTGTTCATTGGATTTTCATTGAAATCACGC	16567
Qy	567 TGTGCTGGATGGTACTTTCGACGTGGCAGCTTTATGACCTCGTGGTCCGAGCTTTGCTGTG-	625
Db	16568 TGTGGAGATGGAGTTCCTTGTGGCATTTTTTCAACTCTTTTGTGCGAGATTTTCAATGC	16627
Qy	626 ---GGTCCACCTCAATTTCTCTCCACCAATTCCTTGAACGAACCAAGGCTCGTAACACTC	682
Db	16628 ACAAGAAACCCGATAACTTGTCTCTCAAGAAATCCCTCCAGTCTTTCCAGCTTGGTTCTCTAA	16687
Qy	683 GAGTCAAGCTCAACTCTCTCAACCATCAGATGCACCCGNACATGCTAAGTCAGCAACCA	742
Db	16588 AGGTACGGT---CCTGTGTACAACCTTCCCTTTCACCTCACCTGTGATGAATTCATTAGCCG	16744
Qy	743 ACGGTGATGTCCC GCCCAAACGTAGACCCACACTCTTCGCGAAAGAGTATTCAAGTTCTCCG	802
Db	16745 ATTCAATCACCGGTC-----TTCAAGGAGAGAATATTCATTCTCGT	16788
Qy	803 AGTTAGCAATTGACAAAATCAAGTCAACAGTCAATGCCAACTCAGGAGAGACGCCATTCT	862
Db	16789 CTGAACAATAAACATCATTTGAATCAAAAGGCTAATGAAGAAAGCAGACAACAACAATTT	16848
Qy	863 CCACATTTCCAATCACTCTCCGACACAGTGTGGCTAGCCGTCACAGCTCGCGGCCCAACTCA	922
Db	16849 CTTCTTTTCAAGCGTTAGCCGCATTTATGTGGAGATGCATTTACAAGACAAGAACAATTTAC	16908
Qy	923 AGCCCGAGGACTACACTGTGTACACTGTGTTTGTGCTATTGTCAGGAAAAAGGTTGATCTCTC	982
Db	16909 CTTATGATCATGAATTCGTTGTAGCCCTTGGCGCTAACAAATGGAACAANAATTTGATCCAC	16968
Qy	983 CAATGCCAGAAAGTTACTTCGGCAACCTTAATTCAGGCAATTTTTCACAGTGACCG---CGG	1039
Db	16969 CATTTGCTTTGAGCTATCTTGTTGTAATTCCTCTCTGCTGTAAAGTCCAAGACCGTTACAT	17028
Qy	1040 CAGGTTTGTACTAGCAAGCCCGATCGAGTTCGCTGGTGGGATGATACACAAGCGGATCG	1099
Db	17029 CCGGTGAGCTACTAGAAAATGATTGGGATGGGCTGCTTTGAAAATGCATGAAGCTGTGA	17088
Qy	1100 TGAAGCATGACGCTAAGGCCATTGATGAAGAAACAAGAGGTGGGAGAGCAACCCGAAGA	1159
Db	17089 TTGGNAACACTAGTGAAGTGGTCTCGGAGACAAATCAAAAACCTGGTTGAATCATCTTATG	17148
Qy	1160 TCCTTTCAGT---ACAAAAGATCCTGGAGTGAACCTGTGTGCTGTTTGGAAAGTTCGCAAGGT	1216

[illegible]

Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (<http://arthur.epm.ornl.gov/pub/xtgrail>), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://genomic.stanford.edu/GENSCAN.html>), and NetPlantGene (<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curations based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins.

Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by repeatmasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to bottom of the chromosome.

We thank the CSLH/WashU/ABI consortium for sequencing BAC clones F6P23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone F13D4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khalak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: at@tigr.org.

```

FEATURES
    source
        Location/Qualifiers
            1..110531
                /organism="Arabidopsis thaliana"
                /cultivar="Columbia"
                /db_xref="taxon:3702"
                /chromosome="II"
            <1..1483
                /note="Sequence from clone T517"
            660..746
                /rpt_family="POLY A"
            join(<770..853,968..1059,1190..1278,1443..1674,1811..2363,
                2453..2556,2790..3035,3146..3272,3336..3579,3664..4244)
                /gene="At2g39850"
            <770..4244
                /gene="At2g39850"
            /note="T517.15"
            join(770..853,968..1059,1190..1278,1443..1674,1811..2363,
                2453..2556,2790..3035,3146..3272,3336..3579,3664..4244)
                /gene="At2g39850"
            /codon_start=1
            /product="putative serine protease"
            /protein_id="AAB95271.2"
            /db_xref="GI:6598792"
            /translation="MATAFHTLLQLLFFFAFAEANDSRKTYLVQMKVGGHRYGSS
                SGQELLEGLVDDSKESFTGFSASLTFRERQKLSKTTIVSSRRLEVSRNRL
                KLOTRSWDFMNLTKAERNPENESDLVAVIDSGIWPYSELFSGSDSPPPGWNKCE
                NITCNKIVGARISYPKKERYKWEKSVIDVTGHGTHVASIVAGRKVKGAGYFLAE
                GTWRGVPNAKIAVYKTCWRVIRKNGRDSVCREDNIIKAIDDAIADKVDIIISYSGF
                QFTPLQDKVSWAFLRALKNGILTSAAAGNYANNKFYTVANGAPWMTVAASLQDR
                IFETKLEGEDKPIIVYDINTVFETQDSYFPLLNKAPPSTRKRELIAERNGYSIL
                SDNDEKDKGVFFFAQINLLDEAKIKHETEEIPREGWPTVAHLSSRGPNQDSFLANIL
                DEQKGLWDYKKDQSKERLAKHETEEIPREGWPTVAHLSSRGPNQDSFLANIL
                KSHMNNCFQDIAAPGLDIAGPENKLSRDPANDYRHLRFNIMSGTSMACPHT
                GLALYKFSFRWSPSAKSLAMTTSMTDDNEFAYGSGHLNATKVRDPLGVYETHY
                QVDYDLXKGVNGTEKLRSVHGSDKIDCSKTEIDHDADLANPTMTARVPLDTPPKK
                VFHRTVYNDGTETLREINRYGDKDFDEIIVDPOLKFSSELGETKFTFTVTGTSK
                RNKNKRAFMRTNLTWLTWEKDSQRVSPIVIYSINGPKACM"
            1484..110330
                /note="Sequence from clone T28M21"
            complement(9924..9995)
                /gene="At2g39860"
            /product="tRNA-Ala"
            /anticodon=(pos:9960..9962,aa:Ala)
            complement(9924..9995)
                /gene="At2g39860"
            /note="T28M21.2"
            complement(11732..11793)
                /rpt_family="{GAAA}n"
            join(13781..14113,14198..14723,14813..14914,15003..15094)
                /gene="At2g39870"
            13781..15094
                /gene="At2g39870"

```

```

/gene="At2g39870"
/note="T28M21.3"
join(13841..14113,14198..14723,14813..14914,15003..15094)
/gene="At2g39870"
/note="unknown protein"
/codon_start=1
/protein_id="AAB95272.1"
/db_xref="GI:2088640"
/translation="MSRLLPHLVSNLPTMEKTTSSLELTTRLGFRNFPYEDPSPF
SPGTSPEDESSDEEDFLAGTRRLAPSTORLSPFKSEKQVAATSPQ
STUSGGSFNSGSRSDLPSPAPTSFRDNDAMDVISAAAGEVARLKLSGVEPHHL
PLQTPESLLRRQNAIIHAELQHQRILIQMLCSAQSRFKLSENRIPRRVNEGLFN
PRVRRNPTWLPQQAAPLKRPSAGTGVFLPRYPDPAAPSDSLKTPVNTPMLQPK
VKQNLNDFETNIVGPRRSQFDYECMLARSTVLARQGNFRAVSGGGLPQDWMY"
complement(15255..15311)
/rpt_family="(TA)n"
15747..15812
/rpt_family="(GAAA)n"
complement(join(<18089..18873,19305..>19623))
/gene="At2g39880"
complement(<18089..>19623)
/gene="At2g39880"
/note="T28M21.4"
complement(join(18089..18873,19305..19623))
/gene="At2g39880"
/codon_start=1
/product="putative MYB family transcription factor"
/protein_id="AAB95273.1"
/db_xref="GI:2088641"
/translation="MNGEISRPPELISRNPKCSFENAIHKAVEAEALAKSDANG
GKSKVGPWLPDEQDEALTRLVKMGCPRNWNLISRGIPRSGKSCRLRWQNLDPILKR
KPSDEEHMIMSAQAVLGNKWSVIAKLLPGRDTNALKNIHNSNLRKPAEOMKIPLL
MSNTEIVQLYPNSVBRISNAKPEHLPOBEETGVLSDKMDDEAKPEPRONSKTG
YRVARMGATSVCKPGYMACPEGLVQASRPDSLAGKFLQSLCYDPIIPKCGGCCN
HQDPTSSVLGSEFVDEEHSAEELDKELISINDLANTAWIRSGKEAQSLKAD
DQFRREYAHKFSGMVNVGVSQMVRODLRLS"
complement(20519..20562)
/rpt_family="(TA)n"
22037..22139
/rpt_family="(TAAA)n"
22753..22816
join(<26166..26265,26637..26861,26940..27034,27116..27305,
27400..27524,27669..27884,27969..28451)
/gene="At2g39890"
<26166..28451
/note="T28M21.5"
join(26166..26265,26637..26861,26940..27034,27116..27305,
27400..27524,27669..27884,27969..28346)
/gene="At2g39890"
/codon_start=1
/product="proline transporter 1"
/protein_id="AAB95274.1"
/db_xref="GI:2088642"
/translation="MTATEAKNRKINVGDDVDVIEIPTAHQISSDSWFOAVFLT
TGNSAVLVGSGTLMVPLGIVGVCLLIATLISLVANTLIJAKLHFEFGRRHRYRD
LAGFIYGRKAYHLTGLQYVNLGFMINGCFILLAGSALKAYVIVLRDHTWKLPFHIAL
ALGICAIFAIGPHLSALGVGVSTLSLIYIVVAIVLSVRDGTWKLDEYIQGSS
LSKLTITGAANLVAFNTGMLPEIQATVQRPVVKNMKALYFQFTAGVLPYAVTF
IGYWAYGSTYLLVNSVNGPLWKALANYSAILOQVSIHLIFASIFYEYMDTKYGIK
GNPFATKNLLFRIMARGGYIAVSTLISALLPELGDMSLTGAVSTFPLTFLANHHY
KAKNNKLNAMQKLHNLNVVFFSLMSVAAAIAAVRLAVDSKNFHFVADL"
join(28838..28973,29224..29414,29559..29658,29750..29787,
29872..29961,30172..30411)
/gene="At2g39900"
28838..30411
/gene="At2g39900"
/note="T28M21.6"
join(29280..29414,29559..29658,29750..29787,29872..29961,
30172..30411)
/gene="At2g39900"
/codon_start=1

```


exon	AYFSFSAVILLTLTLCQSFYAVYAYKPKNSKL"	complement(2238. .2668)	/number=1	gene	/number=4	12843. .13704	/gene="Fl1Cl_40"	complement(join(12843. .12868,13035. .13704))	/note="similarity to predicted protein, Arabidopsis thaliana"
intron				CDS					
exon		complement(2669. .2806)	/number=1						
intron									
exon		complement(2807. .3330)	/number=2						
intron									
exon		complement(3331. .3448)	/number=2						
exon		complement(3449. .4119)	/number=3						
gene	5324. .7543								
CDS	/gene="Fl1Cl_20"	complement(join(5324. .5760,6099. .6562,6678. .7543))	/note="similarity to various predicted genes, Arabidopsis thaliana and Oryza sativa"						
	/codon_start=1								
	/product="putative protein"								
	/protein_id="CAB62297.1"								
	/db_xref="GI:6523029"								
	/translation="METFGTIPPPPLPRLLERRQAPPPQPPPPPPPPPPRL								
	GPRLRLPLPPROQLLRKQPLVLPEPLRPRAIPWLTHDVKKGSYDKLREY								
	VENVRQNESESVLPKSIYEDONLEODTBSKPGINEVVEGSGTKBLEWVISIK								
	DKLQAVNRDRTSGKLCIYKVPHYLHGNDKKSYPQTQVSLGPHYHGROQTQSMECH								
	KRAVNVKRTNGIEVLDAMIELEEKACACYESIVLSSNEFTMLLLDGCFILE								
	LLQVNEGFLKGLYDHPVPFAVSGSHSIQORDIMLEQLPLFVLNLLLEPGTQN								
	QTGLVELVPTFTPLMPTAETLTNSPPRGVSGELHCLDVFHRSLLFPSSGKANYS								
	RVADKHLQVPTVTELRDAGFKPLAKTRDFWDIKFSNGVLEIPGLLIHDGTSKSLFL								
	NLIAPECHIESSNDITPILFMDNLIDSPEDISYLHHCGLIEHSLGSSEVADMPNQ								
	LCQEVPTDRIYLSQLLIEVHRCYKQNSKLSLTKYTLILKYLDPNWAYLSFFAAV								
	ILLIITFSQSYFAAYAFNPSS"								
exon	complement(5324. .5760)		/number=1						
intron									
exon	complement(5761. .6098)		/number=1						
exon	complement(6099. .6562)		/number=2						
intron									
exon	complement(6563. .6677)		/number=2						
exon	complement(6678. .7543)		/number=3						
gene	9545. .11550								
exon	/gene="Fl1Cl_30"	complement(9545. .9615)							
exon			/number=1						
CDS	complement(join(9545. .9615,9706. .9978,10386. .10813,10931. .11550))								
	/note="similarity to various predicted proteins, Arabidopsis thaliana and Oryza sativa"								
	/codon_start=1								
	/product="putative protein"								
	/protein_id="CAB62298.1"								
	/db_xref="GI:6523030"								
	/translation="MAVPSRRSPSPPPPPQPLPFGPKLPPEPHIIRKARNRVQLRP								
	MLTWYFTFMKRLQTRNOQPETREEMVIWIDKMEQVMRDAATTSWDKLCIYRYP								
	LSFKQSKNSYFPHYTVSLGPHHGDHLRPHWYHWRVAVNMVKRTKQIEMYIDAMK								
	ELEERARACYEGPGLSSNKFTQMLVLDGCFVLDLFCGAYEGFSKLGVDNDPVPFAMR								
	GSMSIRDMLEMLNQLPLMPTWPSYKINSQENNNKFFPHIADKKEELHCLDVER								
	RSVLQPSLKPEPRLSRWSKVPVADRQKLLHCVELREAGIKFRRRKSDRLWDIQ								
	FKNGCLEIPKLLIHDGTSKLSLNLIAPFQCHIDSTKQITSIIEVENLINSNEDRVXL								
	QYCGIIHEWLESDEVDLFLKLCQDVVDFPIDSYLPLCLGVDRFYSLIILLITLTCFCE								
	SFFTYYVAFEPNT"								
intron	complement(9616. .9705)		/number=1						
exon	complement(9706. .9978)		/number=2						
intron									
exon	complement(9979. .10385)		/number=2						
exon	complement(10386. .10813)		/number=3						
intron									
exon	complement(10814. .10930)		/number=3						
exon	complement(10931. .11550)								

Query Match

4.3%; Score 68.8; DB 13; Length 111122;

Best Local Similarity 43.9%; Pred. No. 1.3e-05;

Matches 402; Conservative 0; Mismatches 502; Indels 12; Gaps 2;

http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=k21h1
Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <http://complib.ornl.gov/Grail-1.3/>), GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://grain1.zoool.iastate.edu/cgi-bin/sp.cgi>).
Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>).
This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K8A10 and the 3' clone is K3G17.

FEATURES

source

1. 74342
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone="K21h1"
/clone_lib="Mitsui TAC"
complement(join(683..778,1044..1247,1325..1959,2053..2283,2364..2661))
/note="gene_id:K21H1.1
unknown protein"
/codon_start=1
/evidence="not experimental
/protein_id="BAB10939.1"
/db_xref="GI:10177592"
/translation="MNMSDDDDPRGYILRPEELRPMELVRLFSFGDIEKPRSVSDS
EPEHRSFRRHIFVLSVLKLLRFRFSKLLALVGSALFSLNNSFSGLFLRGEV
VMPQTSNYQSFIGHLDTRVSLDMLNREDEKYYAALSIMASKIAYENAAIRKHV
ENHNMKVLGLVDYNEVQEKETQAFIMSTDETTTRNGOETTUVVAFRCETELNSE
DWCSDPDITWELPNIHNGFFMKALQNLNCSWPKEPLSPDRKSLPLAYISIRDSL
KTLIAQNKTFLVLRGSLGALAILFTAVLVIHETELLERIQGVYITQPRVGDGSK
QSEFMKLEKIKYIRFYVNDINVRPLDQDKLMFKHGTCTIYDQNYOAKYMRN
FGDNFFLLRGIKMMWAAILEFIRFTPIAEKSEYSEGWLLKGRGALGIITPVGSN
HTPDQYKNSRVSYNGYEDVPRKDETALKRVIFPTG"

CDS

/note="gene_id:K21H1.1
unknown protein"
/codon_start=1
/evidence="not experimental
/protein_id="BAB10939.1"
/db_xref="GI:10177592"
/translation="MNMSDDDDPRGYILRPEELRPMELVRLFSFGDIEKPRSVSDS
EPEHRSFRRHIFVLSVLKLLRFRFSKLLALVGSALFSLNNSFSGLFLRGEV
VMPQTSNYQSFIGHLDTRVSLDMLNREDEKYYAALSIMASKIAYENAAIRKHV
ENHNMKVLGLVDYNEVQEKETQAFIMSTDETTTRNGOETTUVVAFRCETELNSE
DWCSDPDITWELPNIHNGFFMKALQNLNCSWPKEPLSPDRKSLPLAYISIRDSL
KTLIAQNKTFLVLRGSLGALAILFTAVLVIHETELLERIQGVYITQPRVGDGSK
QSEFMKLEKIKYIRFYVNDINVRPLDQDKLMFKHGTCTIYDQNYOAKYMRN
FGDNFFLLRGIKMMWAAILEFIRFTPIAEKSEYSEGWLLKGRGALGIITPVGSN
HTPDQYKNSRVSYNGYEDVPRKDETALKRVIFPTG"

CDS

/note="emb|CAB62312.1
gene_id:K21H1.2
similar to unknown protein"
/codon_start=1
/evidence="not experimental
/protein_id="BAB10940.1"
/db_xref="GI:10177593"
/translation="MDSDIMMMHMQEKLPEFCNPNSFFSPDRNNYPPFLFNSTHY
QSDHSMTEPGFYSGGLTNPSSIPNTAVSYVFLDKRNNNNNNNTMAAMREMI
FIAYMQTHIDPEAVKPKRNRNYSKDPQSAARHRERISERILQIRLVPQGTIK
MDTASMLDEATHYVFLKQVQLEEQAVVTGGGGGGGGRVLIGGGGTAASGGGGGGG
GVVMGCGCTVGHQWVNAQILR"
9616..10005
/note="gb|AA64534.1
gene_id:K21H1.3
similar to unknown protein"
/codon_start=1
/evidence="not experimental
/protein_id="BAB10940.1"
/db_xref="GI:10177593"
/translation="MDSDIMMMHMQEKLPEFCNPNSFFSPDRNNYPPFLFNSTHY
QSDHSMTEPGFYSGGLTNPSSIPNTAVSYVFLDKRNNNNNNNTMAAMREMI
FIAYMQTHIDPEAVKPKRNRNYSKDPQSAARHRERISERILQIRLVPQGTIK
MDTASMLDEATHYVFLKQVQLEEQAVVTGGGGGGGGRVLIGGGGTAASGGGGGGG
GVVMGCGCTVGHQWVNAQILR"

CDS

/note="gb|AA64534.1
gene_id:K21H1.3
similar to unknown protein"
/codon_start=1
/evidence="not experimental
/protein_id="BAB10941.1"
/db_xref="GI:10177594"
/translation="MAASNLNLLTSLTFLTSIQRSLSLSDNPSTLLPDPGFWPIS
HSDEPDIDGESFVETEDDQVTDRRSLWRRTYYISYGALSANRVPCCPPGSCRYS
YTHNCFRAGPVHPYSRGCSSTTRCR"
14222..15256
/note="gene_id:K21H1.4"
/codon_start=1
/evidence="not experimental
/product="protein kinase-like protein"
/protein_id="BAB10942.1"
similar to unknown protein"

CDS

/note="emb|CAB16798.1
gene_id:K21H1.7
similar to unknown protein"

/db_xref="GI:10177595"
/translation="MEWIRGETIGYGTFTSVLSTRNNDGCEPPLMAVKSADSYGGA
ASLANEKSVLNDLDDCNEIYRCFGEDRTVENGEEMHNLFLYASRGSLESYLKLAG
EVPSESTVRRHTGVSRLGRHIHANGFAHCDLKIENILFGLDGVAKVIFADGLAKRIGD
LTALNYGVOIRGTPLYMAPESVNDNEYSGDVGWALGVVEMFSGKTAWSLKHNF
MSLLRIGVDEVPIMPBELSEQDFLSKCFVKDKRKRAMELLNHPFVTVDDVHD
VLVKEEDFVNNKTEDVSTSPRCPEFPDMVSVSGSQTIDSPDERVASLVTDMIPDW
SVTNSWTVNR"
complement(15607..17817)
/note="gene_id:K21H1.5"
/codon_start=1
/evidence="not experimental
/product="subtilisin-type protease-like"
/protein_id="BAB10943.1"
/db_xref="GI:10177596"
/translation="MGMTVILVFPFVAIVTAETSPYIIMHDSAKPLPFSHRSH
FSTTPTSTVNNKPKIIYAYDSVHGFSALVNSLQRLKHKPGVSVFTKDLVPLKLT
FTSPKIGLNTSGTWPVSNYGAGIVIGIDTWPDSFHDGSGVGVSKWKACE
FNSSSLCNKLLIGAKVFNKGLFANNPDLRETRKIGQYSSPYDTIGHTVHAATAAGNHV
KNASYFSAQGTASGIAPHALAIYKAAWEEGIIYSSDVIAAIDQAIRGVHVHISLSLG
LSFEDDDNUGFLENDPIAVASFAAIQKGVFVVTSGGNDGPPYVSLINGAPIMTVG
AGTGRQFQGLTFEGNRVSVFSPPLFPGEFSPVQPYVIESGSVENETLANRIVCN
ENINIGSKLHQIRFVSGAAVLTIDKLEEDTIKQFPVAFIGSKHRETIIESYASN
KNNATAKLETRKTVIGTAPPEVGTSSRGPTSPQILKPDILAPGTLILSAMPSE
QITGTRALPLFSGFNLLTGTSMAPHVAGVAALIKQVHPNMPSPAISKAITALTLD
NPLVAGCHVSTNKVNLPLGIYDTPQDFINFLCHEAKQSRKLIITRSNTSDACKK
PSPLVNPSTIAYFTSDQSSPKIFKRTLTNVGEAKRSYIVRVGLKLVNVPKLM
FSEKNEKUSYTVRLSEPRGLQNVVIGLVSWDEDEAEFEVSCSVAVTSLVQES"
join(18490..19596,19689..19859,19945..20028,20117..20191,
20278..20342,20515..20605,20698..20883,21116..21279,
21388..21502,21780..21956,22033..22146,22391..22510,
22590..22904,22985..23074,23461..23579,23668..23740,
23831..23887,23939..24091,24221..24343,24422..24519,
24598..24643,24730..24843,24949..25032,25116..25322,
25414..25490,25698..25821,25946..25999,26108..26227,
26298..26372,26460..26600)
/note="gene_id:K21H1.6"
/codon_start=1
/evidence="not experimental
/product="DNA polymerase alpha 1"
/protein_id="BAB10944.1"
/db_xref="GI:10177597"
/translation="MSGDNSTETRRSRGAEASRKDTLERLKAIROGGIRASGGG
YDRIQKPIEDTVDDDEYDALVSRREARGPVVDEGGDGLYLDGEEDWSPSG
PESTDESDDGGFSGRLAKKKKKEQTOQPOVKYNPALKAAATITGBRLSSMTSS
SFKVKETDFAQYEGILDEITAOVTPDESDRKKHRRKLPGLTPVTIFANKALFSVAS
SMGKSESEPTSTYEGDSVMDNLMKEEDMKESEVIPSETMELIGSDIVKEDGSKNI
RTEYKSELGVKEVFTLNATIDMKEKSALSTAGKEMKGVGTENGALLGSSSEGK
TEFDLDADGSLRFFILDAYEEAFGASMTIYLFKGVKMGDTYKSCVVVNTQRCVYA
IPNDSIFPSHELIMLEQEVKDSRLSPESFRGLHEMASKLNEIAQELLQLVNFSM
APVKRNYAFERPDVPAGQYVLKINYSKDRPLDELKGESFALLGSHTSALEHFIL
KRKMGAPWLKISSFTSPSGSWCKFEVTVQSPKDIITLVSEKVVHPVPAVTAI
NLKTLVNEKQNISEIVSALVCFHNKIDVPMAPERKSRGSLSHFTVVRNPEGTYP
IGWKVEKDNKSGNCVLSIENSERALLNRLLELNKLDLILVGHNIFGDLVDLL
ORQAACKVQSSSMWSKIGRLKRSFMPKLSNGSYSGATPGLMSKTLCDTDLCSR
DLLKEVSYSLTDLSTQLNRKETAINDIPKMFOSKTLVELIECGETDAMLWELM
PHLSVLPTLQLTNLSNGLWKTLOGARQRIEYLLHTFHSKKTFLDPKISQRKEI
KSSKRMIDYAPEDRVNVDLADLTLENDSPSGSKTKGPAYAGVLVLEPRGLYDKV
LLLDNFSLYFSDIYQINCFITIPRSEDGVPRLSPSSQTPGILPKLMEHLVSRKSVKL
KMKKETGLKRWELDIQOALKLTANSGLDIEEVKAISKVIOEVNKKYRCLKIDCDGIY
VDLVONHLNLEVIYGDTSIMIHSLDDEEVKAISKVIOEVNKKYRCLKIDCDGIY
KRMILLRKKKYAAVKLQKCKPCEDIERKGVDMVRDWSLLSKELGCLSKILYGG
SCEDVEEAHNEMLKIEEMRNGQVALEKIYITKTLTPPAAYIPDSKVSOPHVQVALRM
RQRCYGFNAKDVTPIYICYBOGNASSASSAGIAERARHPDEKSESRMLVDIYY
LAQQLHPVYVRLCAIQGTSPERLAECGLDPSKYSKSNADTSDPSTSLFATSDSE
ESKPAFTETESDSTFWLKHCPQOEDSTGIIISPAIANOVKQRIQDGFVSMYKQ
IMVCEDESCKHTTRSPRFLLEREGRTVCPNYPNCNGTLLRKYTEADLYKOLSVFCH
ILDTQCSLEKMDVGVRIQVERAMTKIRPAVKSAAAITRSSRDRCAVGMQLTDIYI"
complement(join(26828..27019,27114..27179,27258..27323,
27414..27485,27598..27834))
/note="emb|CAB16798.1
gene_id:K21H1.7
similar to unknown protein"

```

/codon_start=1
/evidence-not_experimental
/protein_id="BAB10945.1"
/db_xref="GI:10177598"
/translation="MGSDVGDRLPPSPSSDLSFLRQILSRTPAQSPSPKSTNV
SSAETFFSVSGVSGVSGTGQDKYAFERHKGAKORNSLRKRDIAQPHNUSE
KKRSKINEMKALQKLIPIKSNKTDKLAEMDEAYELKQLQVQTAVWNGLGNPM
RLQVPPPTTRINETLEQDLNLETLLAAAPHSLEPAKTSQCMCFSTATLL"
complement(join(29326..29586,29960..30082,30390..30512,
30845..30958,31036..31233))
/note="gene_id:K21H1.8
pir|I05432
similar to unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB10946.1"
/db_xref="GI:10177599"
/translation="MDNEONSEIPDPSPFMMFLFGDDMDLWDTYSPVDAADISPPD
TLVNGDGPAPLPGSLRTISWETTFQDSQWNRQSETTSVETDYSSHGNTFT
NEETLERYWRNQLQSTNEQSETGSEETYNASSHGCTFIYEETLEQYWRNWLQSS
TNEQFETSLSEETNPSSHGDIIFYEBLLSTIDTGTGERTGLSEVIDENLRKYEK
RSDDETARCVCOKKLDNEEVSKLGGCHDFHFGCIKNWLMVFNKCPICNREVV"
join(33235..33402,33654..33791,33873..33971,34064..34367,
34837..35225)
/note="contains similarity to MAP3K-like protein kinase
gene_id:K21H1.9"

Query Match 4.2%; Score 67; DB 12; Length 74342;
Best Local Similarity 47.0%; Pred. No. 3.2e-05;
Matches 276; Conservative 0; Mismatches 305; Indels 6; Gaps 2;

QY 817 AAATCAAGTCAACAGTCAATGCCAAGTCAAGGAGAGCGCCATTCACACATTCACATCA 876
Db 37972 AAAGGAAAGCTAAGCGGAGGTTGACTCCGATGATCGGAAATATCGTCTCAACGG 37913

QY 877 CTCCTCCGACACGCTGGCTAGCCGCTACAGCTGGCGGCCCAACTCAAGCCGAGGACTAC 936
Db 37912 GTTTCAGCGTATATGTCGCGATCATGAAACAGTGTCTGAATCCGGAAGAGT 37853

QY 937 ACTGTGTACACTGTGTTGCTGATGACAGGAAAGGTTGATCTCCCAATGCCAAGAGT 996
Db 37852 ATTCATTCGAAGTTACTGCTGGATATGAGAGGAGACTAAACCCCTCCCTTTCAGAAAG 37793

QY 997 TACTTCGGCAACCTAATTCAGGCAATTTTCACAGTACCGCGGAGGTTTGTACTAGCA 1056
Db 37792 TGTTCGGTAACTGCTGGGCTGCGGAGTTCGCGAGTACGACTACTGTGGCGAGATGCTTAC 37733

QY 1057 AGCCGATCGAGTTCGCTGGGATGATACAAAGCGATCGTGAACCATGACCTAAG 1116
Db 37732 AATGGGCTTGTGTTGGGCTGCAATTCGAAATAACAAACCTGTGGTTCACAAACGAATGAA 37673

QY 1117 GCATTGTATGAAGAAACAAAGAGTGAGAGAGCAACCCGAAAGATCTTTCAGTACAAAGAT 1176
Db 37672 GAGTTTAGAGAGTTGCTGAGAAATTTGGGTGAAGAACCGAGTATATGAATGCTTAGGCT 37613

QY 1177 GCTGAGTGAAGTGTGCTGTTGGAAGTTTCGCAAGGTTTCAAGGTTTACGACGTGAT 1236
Db 37612 TTTAG---TAATTCATATAACAATTCCTAGCTCTCCTCGGTTCAATGTCTATGGAACGAT 37556

QY 1237 TTTGATGGGAAACCCAGAGAGTGTGAGGAGTGTTCGACACATAGTTTGTATGGAATG 1296
Db 37555 TTCGGTTGGGTTAAACCAATAGCGGTTTCGAGCTGGACGCGGAAATACCCGATGGTAAA 37496

QY 1297 GTGATTTGTACAAAGGCAAAATGAGGAAAGCAAGCATTTGATGTGGAGATTAGTTTGAA 1356
Db 37495 CTCATTGCTTATCCGG---GATCGAAAGAAACATCGAGTTTCAACAGCTGTTTATCG 37439

QY 1357 GCAATGCTATGGAGAGTTGGAGAAAGATAAAGAGTTTCCTCATGGA 1403
Db 37438 TCTAGTGTGTTGGAAAAATTTATCGACTCATGAAGAGTTTGTGAAGCA 37392

```

RESULT 14

```

AP003279/c
LOCUS
DEFINITION
Oryza sativa chromosome 1 clone P0529E05, HTG 21-FEB-2001
PROGRESS ***, in ordered pieces.
ACCESSION
AP003279.1 GI:13027309
VERSION
HTG; HTGS_PHASE2.
KEYWORDS
Oryza sativa (cultivar:Nipponbare) DNA, clone:P0529E05.
SOURCE
Oryza sativa
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;
Oryza.
REFERENCE
1 (sites)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0529E05
Published Only in DataBase (2001) In press
2 (bases 1 to 137205)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@abrr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
1..137205
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="P0529E05"
BASE COUNT 38429 a 29910 c 31211 g 37505 t 150 others
ORIGIN

Query Match 3.8%; Score 60.8; DB 83; Length 137205;
Best Local Similarity 47.1%; Pred. No. 0.001;
Matches 256; Conservative 0; Mismatches 282; Indels 6; Gaps 2;

QY 860 TCTCCACATTCCAATCACTCTCCGCACACGCTGTGCTAGCGTACACGTCGCGGCCA-- 917
Db 19493 TCTCGTCTCCAGTCACTGTGCGCGAGATATGCTCTCGTGCACGCGCGCGGACG 19434

QY 918 -ACTCAAGCCCGAGGACTACACTGTGTACACTGTGTGTTGTTGATTCAGGAAAGGGTTG 976
Db 19433 GTCTGGCCCGCGAGCGGACGAGCAGTTCCGATGGCGGTGAACCTGCGCGCACAGGCTGC 19374

QY 977 ATCTCCAAATCCAGAGAGTTACTTCGCAACCTAATTCAGGCAATTTTCAGTGCACG 1036
Db 19373 GCCCGGCGATTCCTCCCTGTGTACTTCGCGCAACGCCATCCAGCGCGCAGCAACAGGCA 19314

QY 1037 CGGCAGGTTTGTACTAGCAAGCCCGGATCGAGTTTCGCTGGTGGATGATACAACAAGCA 1096
Db 19313 CGGTGGCGGAGCTGGCGTCCCAACGACCTGCGGTGGCGCGCGCCCAAGCTCAACGCGACC 19254

QY 1097 TCGTGAAGCATGACGCTAAGGCCATTTGATGAAGAAACAAAGAGTGGGAGGACACCGCA 1156
Db 19253 TGGCGGCGTACGACGAGCGCGCGGATACGCGCGCGCGCGGTGGCAGGCGGCTCCCC 19194

QY 1157 AGATCTTTCAGTACAAAGATGCTGGAGTGAACCTGTGTGTTGCTGTTGGAAGTTTCGCAAGGT 1216
Db 19193 GGTGCTTCCCGCTGGGCAACCCGCGCGGTGATCAGATGGGAGCTCGAACCAGT 19134

```

Qy	1217	T---CAAGGTTTACGACGTGGATTTTGGATGGGGAAGCCAGAGAGTGTGAGGAGTGGTT	1273			
Db	19133	TCCCATCTGATGAGGGCAACGACTTCGGGTGGGCGCCGCCCTTCGCGTGGGAGCGGCC	19074			
Qy	1274	CGAACATAGTGTGATGGGAATGCTGTATTGTACCAAGGCAAAAATCGAGGAAGAAGCA	1333			
Db	19073	CGGCACACAAGTTCGATGCGCAAGATGTTCGGGTTTCCCTGCCCGCGCGGACGGCAGCG	19014			
Qy	1334	TTGATGTGGAGATTAGTTTGGGAAGCAAAATGCTAGGAGAGGTTGCGAAGAAATAAAGAGT	1393			
Db	19013	TGGACATCGAGGTGCTCGCCCCGACACCATGCGGCGGCTGCTCCGCACAGCGAGT	18954			
Qy	1394	TCCT	1397			
Db	18953	TCAT	18950			
RESULT 15						
AP003442						
LOCUS	AP003442	151703 bp	HTG			
DEFINITION	Oryza sativa chromosome 1 clone B1096A10, *** SEQUENCING IN					
	PROGRESS ***, in ordered pieces.					
ACCESSION	AP003442					
VERSION	AP003442.1 GI:13486911					
KEYWORDS	HTG; HTGS_PHASE2.					
SOURCE	Oryza sativa (cultivar:Nipponbare) DNA, clone:B1096A10.					
ORGANISM	Oryza sativa					
	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;					
	Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;					
	Oryza.					
REFERENCE	1 (sites)					
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.					
TITLE	Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC					
	clone:B1096A10					
JOURNAL	Published Only in Database (2001) In press					
REFERENCE	2 (bases 1 to 151703)					
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.					
TITLE	Direct Submission					
JOURNAL	Submitted (28-MAR-2001) Takuji Sasaki, National Institute of					
	Agrobiological Resources, Rice Genome Research Program; Kannondai					
	2-1-2, Tsukuba, Ibaraki 305-8602, Japan					
	(E-mail:tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,					
	Tel:81-298-38-7441, Fax:81-298-38-7468)					
COMMENT	NOTE: It currently consists of 1 contigs. Gaps between the contigs					
	are represented as runs of N. The order of the pieces is believed					
	to be correct as given, however the sizes of the gaps between them					
	are based on estimates that have provided by the submitter. This					
	sequence will be replaced by the finished sequence as soon as it is					
	available and the accession number will be preserved.					
	* NOTE: This is a working draft sequence.					
	* This sequence will be replaced					
	* by the finished sequence as soon as it is available and					
	* the accession number will be preserved.					
FEATURES	Location/Qualifiers					
source	1..151703					
	/organism="Oryza sativa"					
	/cultivar="Nipponbare"					
	/db_xref="taxon:4530"					
	/chromosome="1"					
	/clone="B1096A10"					
BASE COUNT	39379 a	36328 c	37093 g			
ORIGIN	50 others					
Query Match 3.88; Score 60.8; DB 83; Length 151703;						
Best Local Similarity 47.18; Pred. No. 0.001;						
Matches 256; Conservative 0; Mismatches 282; Indels 6; Gaps 2;						
Qy	860	TCTCCATTCCTCACTCTCCGACACGCTGTGGCTAGCCGTCACAGTCGCGGCCA--	917			
Db	4748	TCTCGTGTCCAGTCACTGTGCGCGAGATATGGCTCTCGGTGACGCGCGGCGCAGC	4807			

Qy	918	-ACTCAAGCCCGAGGACTACACTGTGTACACTGTGTTTGGCTGTATTCAGGAAAGAGGTTG	976
Db	4808	GTCTGGCGCCGACGACGAGCAGCTTCCGATGGCGGTGAACCTGGCGGCACAGCTGC	4867
Qy	977	ATCCTCCAATGCCAGAAACTTACTTTCGGCAACTAATTCAGCAATTTTCACAGTGAACG	1036
Db	4868	GCCCGGGGATCTCCCTGTGTACTTTCGGCAACGCCATCCAGAGCGCAGCAACACGGCCA	4927
Qy	1037	CGCAGGTTTGTACTAGCAAGCCGATCGAGTTCGCTGGTGGATGATACAACAAGCGA	1096
Db	4928	CGTGGCGGAGCTGGCTCAAGGACCTCGGTGGCGCGGCCCAAGCTCAACCCGAGCC	4987
Qy	1097	TCGTGAAGCATGACGCTAAGGCCATTGTATGAAGAAACAAGGAGTGGGAGAGCAACCCGA	1156
Db	4988	TGGCGGCTACGACGACGGCGCATACGCCGCCGCGCGCTGGCAGGCGCTCCCC	5047
Qy	1157	AGATCTTTCAGTACAAGATGCTGGAGTGAACCTGTGTGCTGTGTGGAAAGTTTCGCCAAGT	1216
Db	5048	GGTGCTTCCCGCTTGGCGAAACCGCGAGCGCGGTGATCACGATGGGAGCTTCGAACCGGT	5107
Qy	1217	T---CAAGGTTTACGACGTGATTTTGGATGGGAAAGCCAGAGAGTGTGAGGAGTGGTT	1273
Db	5108	TCCCCATGTATAGGGCAACAGATTCGGGTGGGCGGCCCTCGCGTGGCGAGCGGCC	5167
Qy	1274	CGAACATAGGTTTGATGGAATGGTGTATTTCTACCAAGGCAAAAATGGAGGAAGAAGCA	1333
Db	5168	CGGCAACAAGTTCGATGGCAAGATGTGGCGTTCCTTGCCTGGCGCGGGACCGCAGCG	5227
Qy	1334	TTGATGTGGAGATTAGTTTGGAGCAAAATGCTATTCGGAGAGGTTGGAGAAAGATAAAGAT	1393
Db	5228	TGGACATCGAGGTGCTCTCGCCCCGACACCATGCGCGCTGCTCCGCGCAGACGAGT	5287
Qy	1394	TCCT 1397	
Db	5288	TCAT 5291	

Search completed: November 5, 2001, 16:46:27
Job time: 10557 sec

Search completed: November 5, 2001, 16:46:27
Job time: 10557 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:00:39 ; Search time 5816.79 Seconds
(without alignments)
2635.909 Million cell updates/sec

Title: US-08-894-356C-2
Perfect score: 1622
Sequence: 1 GAACCATTCGAATCCCAATTAA.....AAAAAAAAAAAAAAAAAAAA 1622

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*
12: gb_est12.*
13: gb_est13.*
14: gb_est14.*
15: gb_est15.*
16: gb_est16.*
17: gb_est17.*
18: gb_est18.*
19: gb_est19.*
20: gb_est20.*
21: gb_est21.*
22: gb_est22.*
23: gb_est23.*
24: gb_est24.*
25: gb_est25.*
26: gb_est26.*
27: gb_est27.*
28: gb_est28.*
29: gb_est29.*
30: gb_est30.*
31: gb_est31.*
32: gb_est32.*
33: gb_est33.*
34: gb_est34.*
35: gb_est35.*
36: gb_est36.*
37: gb_est37.*
38: gb_est38.*
39: gb_est39.*
40: gb_est40.*
41: gb_est41.*
42: gb_est42.*
43: gb_est43.*
44: gb_est44.*
45: gb_est45.*
46: gb_est46.*
47: gb_est47.*

44: em_esthum10.*
45: em_esthum11.*
46: em_esthum12.*
47: em_esthum13.*
48: em_esthum14.*
49: em_esthum15.*
50: em_esthum16.*
51: em_esthum17.*
52: em_esthum18.*
53: em_esthum19.*
54: em_esthum20.*
55: em_esthum21.*
56: em_esthum22.*
57: em_esthum23.*
58: em_esthum24.*
59: em_esthum25.*
60: em_esthum26.*
61: em_esthum27.*
62: em_esthum28.*
63: em_estin1.*
64: em_estin2.*
65: em_estin3.*
66: em_estin4.*
67: em_estin5.*
68: em_estom1.*
69: em_estom2.*
70: em_estov1.*
71: em_estov2.*
72: em_estpl1.*
73: em_estpl2.*
74: em_estpl3.*
75: em_estpl4.*
76: em_estpl5.*
77: em_estpl6.*
78: em_estpl7.*
79: em_estpl8.*
80: em_estpl9.*
81: em_estpl10.*
82: em_estro1.*
83: em_estro2.*
84: em_estro3.*
85: em_estro4.*
86: em_estro5.*
87: em_estro6.*
88: em_estro7.*
89: em_estro8.*
90: em_estro9.*
91: em_estro10.*
92: em_estro11.*
93: em_estro12.*
94: em_estro13.*
95: em_estro14.*
96: em_estro15.*
97: em_estro16.*
98: em_estro17.*
99: em_estro18.*
100: em_estro19.*
101: em_estro20.*
102: gb_est25.*
103: gb_est26.*
104: gb_est27.*
105: gb_est28.*
106: gb_est29.*
107: gb_est30.*
108: gb_est31.*
109: gb_est32.*
110: gb_est33.*
111: gb_est34.*
112: gb_est35.*
113: gb_est36.*
114: gb_est37.*
115: gb_est38.*
116: gb_est39.*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description
1	152	9.4	583	113	AW221050	AW221050 EST297519
2	151.4	9.3	690	155	BG599447	BG599447 EST504342
3	150.6	9.3	606	113	AW221049	AW221049 EST297518
4	144.6	8.9	570	142	BE922784	BE922784 EST426553
5	140.6	8.7	763	155	BG591758	BG591758 EST499600
6	133.4	8.2	591	167	BE434257	BE434257 EST405335
7	128.6	7.9	487	118	AW616206	AW616206 EST037245
8	125.2	7.7	497	122	AW930934	AW930934 EST356777
9	124	7.6	482	144	BF096979	BF096979 EST400516
10	119.2	7.3	518	119	AW650280	AW650280 EST328734
11	114.6	7.1	564	142	BE921494	BE921494 EST425179
12	114.6	7.1	659	155	BG598154	BG598154 EST496832
13	112.4	6.9	508	173	BG097054	BG097054 EST461573
14	108.8	6.7	574	142	BE922746	BE922746 EST426515
15	103.6	6.4	403	143	BF053202	BF053202 EST438432
16	103.6	6.4	788	164	BE187619	BE187619 EST336180
17	89.4	5.5	687	111	AW102336	AW102336 sdb6d06.y
18	85.8	5.3	575	142	BE920468	BE920468 EST424237
19	79.8	4.9	549	21	A1490445	A1490445 EST248771
20	77.8	4.8	783	155	BG584708	BG584708 EST486469
21	75.2	4.6	451	167	BE433920	BE433920 EST404998
22	75.2	4.6	578	143	BF051101	BF051101 EST436276
23	75	4.6	542	167	BE436185	BE436185 EST407263
24	73	4.5	411	142	BE922378	BE922378 EST426147
25	72.8	4.5	456	142	BE923572	BE923572 EST426147
26	72.6	4.5	354	142	BE922950	BE922950 EST426719
27	71.6	4.4	414	142	BE921952	BE921952 EST425721
28	68.8	4.2	368	142	BE921953	BE921953 EST425722
29	67.2	4.1	603	118	AW616119	AW616119 EST296884
30	65.2	4.0	386	173	BG097680	BG097680 EST462199
31	64.6	4.0	686	120	AW776512	AW776512 EST335577
32	62.2	3.8	642	153	BG448305	BG448305 NF070A05E
33	61.8	3.8	676	153	BG455847	BG455847 NF071A12P
34	61.8	3.8	681	151	BF639071	BF639071 NF079A08P
35	59.2	3.6	694	117	AW506086	AW506086 EST315734
36	59.2	3.6	829	155	BG584678	BG584678 EST486439
37	58.8	3.6	411	142	BE921804	BE921804 EST425573
38	58.6	3.6	658	146	BF278256	BF278256 GA_EB003
39	57.8	3.6	526	105	AL386584	AL386584 MIBC35E10
40	57.8	3.6	649	151	BF643475	BF643475 NF004H05E
41	56.6	3.5	708	138	BE660849	BE660849 7-C8 Gmax
42	56.4	3.5	311	166	BE353747	BE353747 EST355090
43	55.8	3.4	307	142	BE920010	BE920010 EST423779
44	55.8	3.4	489	138	BE660848	BE660848 330 GmaxS
45	54.2	3.3	648	154	BG457806	BG457806 NF034C06P

ALIGNMENTS

RESULT	1
AW221050	
LOCUS	AW221050 583 bp mRNA EST 07-DEC-1999
DEFINITION	EST297519 tomato fruit mature green, TAMU Lycopersicon esculentum
ACCESSION	cdna clone cLEF3F3, mRNA sequence.
VERSION	AW221050
KEYWORDS	AW221050.1 GI:6532734
SOURCE	EST.
ORGANISM	tomato.
REFERENCE	Lycopersicon esculentum
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
	1 (bases 1 to 583)
	Alcala,J., Vrabalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.

TITLE	Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
JOURNAL	Generation of ESTs from tomato fruit tissue
COMMENT	Unpublished (1999)
	Contact: David Frisch
	Clemson University Genomics Institute
	Clemson University
	100 Jordan Hall, Clemson, SC 29634, USA
	Tel: 864 656 4366
	Fax: 864 656 4293
	Email: dfrisch@CLEMSON.EDU
	5 prime sequence.
FEATURES	
source	Location/Qualifiers
	1..583
	/organism="Lycopersicon esculentum"
	/cultivar="TA496"
	/db_xref="taxon:4081"
	/clone="cLEF3F3"
	/library="tomato fruit mature green, TAMU"
	/tissue_type="fruit pericarp"
	/dev_stage="mature green (3-5 days pre-ripening)"
	/lab_host="SOLR"
	/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2: XhoI; cLEF3: Fruit were tagged at the 1cm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"
BASE COUNT	142 a 140 c 117 g 184 t
ORIGIN	
	Query Match 9.4%; Score 152; DB 113; Length 583;
	Best Local Similarity 56.4%; Pred. No. 1.3e-27;
	Matches 325; Conservative 0; Mismatches 245; Indels 6; Gaps 2;
QY 58	CAAAGTCTTGGAGAAATGCGTGTGGCCACACCGGAGCCGTCGCCAGTTTACAGT 117
DB 9	CACCGTGATTGAGCAATGTCAAGTTGCGCCACCTCCCGCGGCGCAACGAGGTGATAC 68
QY 118	CCACATGTCTGTTTTCGACATGCGTGTGATCTCTGATGCAGACACCATCTGCATTT 177
DB 69	CCCTCTTACTATTATTGACCATGTTGGTTAGGGTTTCGCCGTATGAGCGGATATATT 128
QY 178	CTACAGATTCCGCGCATCTTCTCCCAACTC---TAAATTTATCATTTTCATCCATTAATC 234
DB 129	TTACAGCTCTCCATTTTCAACCCGATTTTCGTTCAAAACATTTATCTCTCTTAAAA 188
QY 235	GTCCCTTTCCCTTGTCTCTCAACACATTTCTTCGTTAGCGGGAATTTGATTTGGCCGT 294
DB 189	TTCACTCTCCCTCCTCCTCAACACATATAGCGCTTAGCGGAAACGTTGCTTGTCCACT 248
QY 295	AGATTCTCCGATAGATGCGCGAGTTGCGTTACAGAAAGGGGACCTCCCTTTCTTAAC 354
DB 249	AGATAAAACG---GATATCTGAGTTACGTTATGTGACAGGAGATTTCTGTGTGTTAC 305
QY 355	AATTGAGAAATCGAGCATGATTTTGTATTTCTCGCGGAGATCATCAGAGGATTTCTTA 414
DB 306	TTTTTTCGAGACTGATATGAATTTCAATATCTCATTTGTTGACCATCCGCGTAAGGCTAA 365
QY 415	TAAATTAACGATTTGATTCGCGAGTGCAGCGATTCACCAACCGATTTAACCTCCGCGGACGAAGT 474
DB 366	GGATTTTATCACTTTGTTCTTAAGTTAGGGAACCACTAAGGATGCACCGGGGTCCAACT 425
QY 475	ATTACCACTTTTTCCTTACAGGTGACGGTGTCTCAACACCGGTATATGATTCGACG 534
DB 426	AGCCCGCTCTTACGCACTTTCAGGTGACACTTTTTCGGAATCTTTGGTGTATCCATTTGGTTT 485
QY 535	CAATCTTCATCAAGTCTTCTGTCGACCTTCTTTCTGCAATTTAAATAAATATGCGGT 594
DB 486	CATTAACCATCATGTTGTTGTCGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 545
QY 595	TTTGGTTGACAAATCCAAATGAGATTTCAATTAAGTT 630

```
Db 546 TCTACTCCACAAATTACGTGGACATGACCAATCTT 581
RESULT 2
Bg599447
LOCUS Bg599447 690 bp mRNA EST 12-APR-2001
DEFINITION EST504342 cSTS Solanum tuberosum cDNA clone cSTS5122 5' sequence,
mRNA sequence.
ACCESSION Bg599447
VERSION Bg599447.1 GI:13616583
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 690)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemiango,A.,
Bougri,O., Bueli,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.
FEATURES
Location/Qualifiers
source 1..690
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS5122"
/clone_lib="cSTS"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
BASE COUNT 178 a 156 c 144 g 212 t
ORIGIN
Query Match 9.3%; Score 151.4; DB 155; Length 690;
Best Local Similarity 57.1%; Pred. No. 1.9e-27;
Matches 316; Conservative 0; Mismatches 231; Indels 6; Gaps 2;
Qy 70 GAAATGCCGTGTGGCCACCACCGGACGCGGTGCGGAGTTTACAGTCCCACTGTGCTT 129
Db 2 GCAATGCAAGTTGGCCACCTCCGGGGGGCGAGGAGTGACACTCCCTCTACTTA 61
Qy 130 TTTGCACATGGGATGGTGTCTCTGATGACAGAACACATCGATTTCTACAGATCCG 189
Db 62 TTTTGATCATGTTGGTTAGGTGCTCCACCGATACGGCGGATATTTCTACAAGTCCC 121
Qy 190 CCATCCCTGTCCCAACTC--TAATATTATCATTTTCATCCATTAATCGTCCCTTCCCT 246
Db 122 CATTTCAAAACCGATTTGCTTCAAAACATTTATTCCTCTTAAATTCACCTCCCT 181
Qy 247 TGTTCCTCAACACATTTCTTCGTTAGCGGGAATTTGATTGGCGGTAGATTCTCCCGA 306
Db 182 CACTCTCAAAACACTATAGCCCTTAGTGGAAACGTTGCTCGTCCACTAGATCAACG- 240
Qy 307 TAGAATGCCGAGTTGGTTACAGAAGGGAGCTCCGTTTCTTTTAAACAATTCGAATC 366
Db 241 --GATATCCTGAGTTACATTATGTGACAGGAGATCTGTATCTGTTACTTTTCTGAGC 298
Qy 367 GAGCATGGATTTTCATATCTCCGCGGAGATCATCAGAGGATTTCTTATAAATTCACGA 426
Db 299 TGATATGGAATTCATATCTCATTTGTTGACCATCCCGGTAATGCTAAGGATTTTATCA 358
```

```
Qy 427 TTTGATTCGCCAGCTGCCAGAACCGATTGTAACTCTCGGCGACGAAGTATTACCACITTT 486
Db 359 CTTTGTTCCTACGTAGGGGAACCTAAGGATGCACCCGGGGTCCCAATTAGCTCCGGTCT 418
Qy 487 TGTTCACAGTGAGGGGTGTTCTCCACACCGGATATATGCAATGGACGCAATCTTCATCA 546
Db 419 AGCCATTCAAGTGACACTTTTTCGGAATCTTGGTGTATCCATCTGTTTCACTAACCATCA 478
Qy 547 AGTCTTGGTGATGCCAGTCTCTTTCTGCATTTTAATAAATATATGGGTTTGGTTGACAA 606
Db 479 TGTGTTGGTGATGGAGCTACCATAGTAGGTTCTATTAGGCGGTGGGCTCTACTCCATAA 538
Qy 607 ATCCAAATGGAGAT 619
Db 539 ATTCGGTGGAGAT 551
RESULT 3
AW221049
LOCUS AW221049 606 bp mRNA EST 07-DEC-1999
DEFINITION EST297518 tomato fruit mature green, TAMU Lycopersicon esculentum
cDNA clone cLEF3F1, mRNA sequence.
ACCESSION AW221049
VERSION AW221049.1 GI:6532733
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 606)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ann,S., Ronning,C.M.,
Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
Location/Qualifiers
source 1..606
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF3F1"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
BASE COUNT 147 a 143 c 123 g 193 t
ORIGIN
Query Match 9.3%; Score 150.6; DB 113; Length 606;
Best Local Similarity 56.6%; Pred. No. 2.9e-27;
Matches 320; Conservative 0; Mismatches 239; Indels 6; Gaps 2;
Qy 58 CAAAGTTCTTGAGAAATGCCGTGTGGCCACCGGACCGGCGGAGTTTACAGT 117
Db 9 CACCGTGATTGAGCAATGTCAAGTTGGCCACCTCCCGGGCGGCAACGAGGTGATACT 68
```

```
QY 118 CCCACTGCTCTTTTCGACATCGGATGTTGATCTCTGATGCAGAACACCATCTGCATTT 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69 CCCTCTTACTTATTTTGACCATCTTTGGTTAGGTTTCGCCGCTATGAGCGGATATATT 128
QY 178 CTACAGATTCGCCATCCTTGTGCCAACTC---TAATTTATCATTTTCATCCATTAATC 234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 129 TTACAAGCTCTCCATCTTTTCAAAACCGATTTTCGTTTCAAAACATTTATTCCTCTCTTAAAA 188
QY 235 GTCCTTTCCCTTGTCTCAAAACACTTTCTCCGTTAGCCGGGAATTTGATTTGGCGGCT 294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 TTCACTCTCCCTCACTCTCAAAACACTATAGCCCTTAGCCGGAACGTTCTGTCTCCACT 248
QY 295 AGATTCTCCGATAGATGCGGAGTTGCGTTTACAAGAAAGGGACTCCGTTTCTTTTAAAC 354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 249 AGATACAAACG---GATATCCTGAGTTACGTTATGTGACAGGAGATCTCTGTCTGTAC 305
QY 355 AATTGCAATCAGCATGATTTTGAATATCTCGCGGGAGATCATCAGAGGATTTCTTA 414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 TTTTTCGAGACTGATGAAATTTCAATTTATCTCATTTGTTGACCATCGCGGTAAGGCTAA 365
QY 415 TAAATCAACGATTTGATTCGCGAGCTGCCAGAACCGATTGTAACTCCCGCGACGAAGT 474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 GGATTTTATCACTTTGTTCTTAAGTTAGGGGAACCTAAGGATGCACCGGGGTCCAAC 425
QY 475 ATTACCACTTTTTCGCTTTACAGGTGACGGTGTCTTCAACACCGGTATATGCATTTGGAGC 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 AGCCCGCTCTTAGCCATTCAGGTGACACTTTTTCGGAATCTTTGGTGTATCCATTTGTTT 485
QY 535 CAATCTTCATCAAGTCTTTCGTTGATGCCAGCTCTTTTCTGCATTTTAAATTAATTTGGT 594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 486 CACTAACCATCATGTTGTTGCTGATGAGCTACTATAGCAGGGTTTCAATTAAGCGGTGGC 545
QY 595 TTTGGTTGCAAAATCCAATCGAGAT 619
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 546 TCTACTCCACAAATTCGGTGACAT 570

RESULT 4
BE922784 570 bp mRNA EST 02-OCT-2000
LOCUS EST426553 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION C5TB22E2 5' sequence, mRNA sequence.
ACCESSION BE922784
VERSION BE922784.1 GI:10448860
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 570)
AUTHORS van der Hoeven,K.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.

FEATURES
source 1..570
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="C5TB22E2"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University)."

```

```
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen.
BASE COUNT 139 a 136 c 115 g 180 t
ORIGIN

Query Match 8.9%; Score 144.6; DB 142; Length 570;
Best Local Similarity 57.0%; Pred. No. 9.2e-26;
Matches 305; Conservative 0; Mismatches 224; Indels 6; Gaps 2;

QY 66 TTGAGAAATCCCGTTGCGCCACCAACCGAGCGCGTTCGCGAGTTTACAGTCCCACTGT 125
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 37 TAGAGCAATGCTCAAGTTGCGCCACCTCCGGCGCGCAGCGAGGTGCACATCCCTCTTA 96
QY 126 CGTTTTTCGACATGCGATGTTGTTGATCTCTGATGCAGAACACCATCTGCATTTCTACAGAT 185
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 97 CTTATTTTGATCATGTTTGGTTAGGTTCCACCGGTATACGGCGGATATTATTCTACAAGC 156
QY 186 TCCGCCATCCTGTCCCAACTC---TAAATTTATCATTTTCATTTAAATCGTCCCTTT 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 157 TCCCATTTCCAAACCGGATTTCTGTTCAAAACATATTCTCTCTTTAAAAATTCACCTCT 216
QY 243 CCCTGTGTTCTCAACACTTTTCTCCGTTAGCGGGAATTTGATTGGCCGCTAGATTCCT 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 CCCTCACTCTCAAAACACTATACGCCCTTAGCTGGAACGTTGCTCGTCCACTAGATAACA 276
QY 303 CCGATAGATGCGCGGAGTTGCGTTTACAAGAAAGGGACTCCGTTTCTTTTAACAATTCGAG 362
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 ACG---GATATCCTGAGTTACATTTATGTGACAGGAGATTCTGTATCTGTACTTTTCTG 333
QY 363 AATCGACATGATTTTGATTATCTCCCGGAGATCATCAGAGGATTTCTTATAAATTCA 422
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 AGACTGATATGGATTTCATATCTCATTTGTTGACCATCCGCTTAATGCTAAGGATTTT 393
QY 423 ACGATTTGATTTCCGACGCTGCCAGAACCGATTGTAACTCCCGCGAGAAAGTATTACCAC 482
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 ATCACTTTGTTTCTACGTTAGGGAACCTAAGGATGCACCGGGTCCAAATTAGCTCCGG 453
QY 483 TTTTGTGTTTACAGGTGACGGTGTCTTCCCAACACCGGTATATGCATTTGGAGCAATCTTC 542
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 454 TCTTAGCCATTCAGTGACACACTTTTCCGGAATCTTGTGTATCCATTTGGTTTCACTAAC 513
QY 543 ATCAAGTCTTGTGTGATGCCAGTTCTTTTCTGCATTTTAAATTAATTTAGGTTT 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 514 ATCATGTTGTTGTTGATGGAGCTACCATAGTAGGTTTCATAGGCGCTGGGCTCT 568

RESULT 5
BG591758 763 bp mRNA EST 12-APR-2001
LOCUS EST499600 P. infestans-challenged leaf Solanum tuberosum cDNA clone
DEFINITION BPL110A14 5' sequence, mRNA sequence.
ACCESSION BG591758
VERSION BG591758.1 GI:13609898
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 763)
AUTHORS Zhang,P., Hernandez,M., Tornqvist,C.-E., Wirtz,U., Loukolanov,A.,
Rangel,P., Haberland,G.T., Cho,J., Chiemingo,A., Bougri,O., Buell
,C.R., Ronning,C.M., Helgeson,J. and Baker,B.
TITLE Generation of ESTs from Potato Leaves Challenged with Phytophthora
JOURNAL Infestans, Incompatible Reaction
COMMENT Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com

```

```
FEATURES                               Seq primer: M13F-R.
source                                Location/Qualifiers
1..763
  /organism="Solanum tuberosum"
  /cultivar="Kennebec"
  /db_xref="taxon:4113"
  /clone="BP110A14"
  /clone_lib="P. infestans-challenged leaf"
  /tissue_type="leaf"
  /dev_stage="6 week old"
  /lab_host="SOLR"
  /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Whole plants were challenged with 450,000
sporangia/ml P. infestans US-1(US 940501) in Biotron
(Madison, Wisconsin). Leaf tissue was collected at 1, 2,
5, 12, and 24 hours post-challenge and frozen in liquid
nitrogen immediately upon removal. Kennebec plants showed
no signs of HR. Katahdin plants (susceptible to P.
infestans US-1) were used as controls and showed
infection. NOTE: We cannot exclude the possibility that
this sequence is actually derived from Phytophthora rather
than potato."
BASE COUNT      212 a 162 c 156 g 233 t
ORIGIN

Query Match      8.7%; Score 140.6; DB 155; Length 763;
Best Local Similarity 56.3%; Pred. No. 9,8e-25;
Matches 346; Conservative 0; Mismatches 259; Indels 10; Gaps 4;

QY 75 GCGGTGTCGCCACACCGGACGGCGTC -GCCGAGTTTACAGTCCCACTGTCGTTTTTC 133
Db 2  GTCAAGTTGCGCACCTCGCGGGCGCGCAGCGAGGTGACACTCCCTCTTACTTATTTT 61

QY 134 GACATCGGATGTTGATCTCTGATCGACAGACACCATCTGCTATTTCTACAGATTCGCCCAT 193
Db 62 GATCATGTTTGGTTAGGTTTCCACCGTATACCGCGGATATTTCTCAAGTCCCCCAT 121

QY 194 CCTGTGTCCAA---CTCTAAATTTATCAATTCATCAATTAATTCGTCCTTTCCCTTGT 250
Db 122 TCCAAACCGGATTTGTTCCAAACATTTATCTACTCTTTAAATTAATCACTCTCCCTCACT 181

QY 251 CTCAAACACTTTCTTCGTTAGCGGGGATTTGANTTGGCCGGTAGATCTCCCGATAGA 310
Db 182 CTCAAACACTACACGCCCTTAGCAGGAATGTTGTCGTCCTCACTAGAT---ACAAATGGA 238

QY 311 ATGCCGGAGTTGGTTTACAAGAAAGGGGACTCCGTTCTTTTAAACAATTCAGAATCGAGC 370
Db 239 TATCTGAGTTAGGTTATCTGACAGGAATTTCTGTATCTGTACTTTTTTTGAGACTGAT 298

QY 371 ATGGATTTTGTATATCTCCCGGAGATCATCAGAGGATTTCTTATAATTCACGATTTG 430
Db 299 ATGGATTTCAATATATCTATCTGTTGACCATCCCGGCTAATGCTAAGGATTTTATCACTTT 358

QY 431 ATTCGCCAGCTGCCAAGACCGATTTGAACCTCCGGGAGCAAGTATATACCACCTTTTGGT 490
Db 359 GTTCTCAAGTTAGGGAACCTAAGGATGACCGCGGGTCCAAATTAGCTCCGGTCTTAGCC 418

QY 491 TTACAGGTGACCGGTGTTCTCCCAACACCGGTATATGATTCATTCAGCGCAATCTCATCAAGTT 550
Db 419 ATTCAGTGACACTTTTTCGGAATCTTGGTGATCTCATCTGTTTTCACTAACCATCATGTT 478

QY 551 CTTGGTGATGCCAGTCTTTTCTGCAATTTTAAATAATATGGGTTTTGGTTGACAATCC 610
Db 479 GTTGGTGATGGAGCTACCATAGTAGGGTTTCATTAAGCGGTGGGCTCTACTCCATAAATTC 538

QY 611 AATGGAGATTCATTAAGTTCCTTCCACTTTCTCTACCTATGTACGACAGATCTGTC 670
Db 539 GTGGGAGAT---GAACAATTTATTCGAATGAGCTAAATTTCCATTTTATGATAGGTCGCTA 595

QY 671 GTGCAAGATCCATTT 685
Db 596 GTAAGGACCCATAT 610
```

```
RESULT 6
BE434257
LOCUS BE434257 591 bp mRNA EST 24-JUL-2000
DEFINITION EST405335 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG15H6, mRNA sequence.
ACCESSION BE434257
VERSION BE434257.1 GI:9432100
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 591)
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Niernan,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
TITLE Generation of ESTs from tomato fruit tissue, breaker stage
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
FEATURES
source
1..591
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG15H6"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKMCluadapt; Site_1: EcoRI;
Site_2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."
BASE COUNT 159 a 129 c 121 g 182 t
ORIGIN
```

```
Query Match      8.2%; Score 133.4; DB 167; Length 591;
Best Local Similarity 56.6%; Pred. No. 6e-23;
Matches 289; Conservative 0; Mismatches 216; Indels 6; Gaps 2;

QY 208 TAAATTTATCATTTTCATCCATTAATCGTCCTTTCCCTTTGTTCTCAACACATTTCTTCC 267
Db 57 TCAAAACATATTTCCTCTCTTAAAAATTCACCTCTCCCTCACTCTCAACACATATACGCC 116

QY 268 GTTAGCCGGGAATTTGATTGGCCGGTAGATTCTCTCCGATAGATTCGCCGAGTTGCGTTA 327
Db 117 CTTAGCCGGGAACAGTTGCTGTCCACTAGATACAAACG---GATATCCGAGTTACGTTA 173

QY 328 CAAGAAAGGGGACTCCGTTTCTTTAAACAATTCAGAATTCAGCATGGATTTTGTATATCT 387
Db 174 TGTGACAGGAGATTTCTGTCTCTGTTACTTTTTTCGAGACTGATATGAATTTCAATATCT 233

QY 388 CCGCCGAGATCATCAGAGGATTCCTTATAAATTCACGATTTGATTCGCCAGCTGCCACA 447
Db 234 CATTTGTCACCATCCGCTAAGGCTAAGGATTTTATCACTTTTGTCTTAAGTTAGGGGA 293

QY 448 ACCGATTTGAACCTCCGGCGGAGAGTATTACACATTTTGTCTTACAGGTGACGGTGT 507
Db 294 ACCTAAGATGACACCGGGGTCCAACTAGCCCCGCTCTTAGCCATTCAGTGCACATTTT 353
```

QY	508	CTCCAACACCGGTATATGTCATCTGGACGCAATCTTCATCAAGTCTCTTGGTGATGCCAGTTC	567
Db	354	TCCGAATCTTGGTGATATGTCATCTGGTTCACCTAACCATCATGCTTCTGGTGATGAGCTAC	413
QY	568	TTTTCTGCATTTTAAATAATATGCGGTTTGGTTTGACAAATCCCAATCGAGATTCATTAAA	627
Db	414	TATAGCAGGGTTCATTAAAGCGGTGGGCTCTACTCCCAAAATTCGGTGACAT---	GAACA 470
QY	628	GTTCCCTTCCACTTTCTCTACCTATGACGACAGATCTGTGGTGCAAGATCCATTCA	687
Db	471	ATTCTTATCGAATGAGCTAAATTCCTATTTTATGATAGTCCGTACTAAAGACCCATATGG	530
QY	688	TATTTCGTGAAAATCTACAAATGAAGAAAA	718
Db	531	ACAAGGATGTCCCATCTGGGAAGAAATGAAA	561

RESULT 7

AW616206

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

EST307245 L. hirsutum trichome, Cornell University Lycopersicon

hirsutum cDNA clone cLHTD15 5', mRNA sequence.

AW616206

AW616206.1 GI:7322240

EST.

Lycopersicon hirsutum.

Lycopersicon hirsutum.

Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 487)

van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.

Generation of ESTs from wild tomato (Lycopersicon hirsutum) trichomes

Unpublished (2000)

Contact: David Frisch

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 4366

Fax: 864 656 4293

Email: dfrisch@CLEMSON.EDU

5 prime sequence.

FEATURES

source

Location/Qualifiers

1..487

/organism="Lycopersicon hirsutum"

/db_xref="taxon:62890"

/clone="cLHTD15"

/clone_1lb="L. hirsutum trichome, Cornell University"

/tissue_type="trichome"

/dev_stage="mixed stages"

/notes="Leaves of various stages were shaken in liquid nitrogen, shearing off trichomes. This procedure yielded a mixture of cells which is highly enriched for trichome likely with minor contaminations of other types of leaf cells"

BASE COUNT

ORIGIN

116 a

117 c

91 g

163 t

Query Match

Best Local Similarity

Matches

274; Conservative

7.98; Score 128.6; DB 118; Length 487;

57.28; Pred. No. 9.4e-22;

0; Mismatches 199; Indels 6; Gaps 2;

QY	85	GCCACCGGACCGCGTCGCCGAGTTTACAGTCCCACTGCTCGTTTTTCGACATGCCGATG	144
Db	1	GCCACCTCCCGACGGCGAACGGAGGTGGTACTTCCCTTACTTATTTTGACCATGTTT	60
QY	145	GTTGATCTCTGATCGAAGAACCACTATCTGCATTTCTACAGATTTCGCCCATCTCTTGCCAA	204


```
QY 595 TTGTTGACAAA 607
||| | |||||
Db 552 TTTACTCAACAA 564

RESULT 12
BG598154
LOCUS EST 659 bp mRNA 12-APR-2001
DEFINITION EST496832 cSTS Solanum tuberosum cDNA clone cST520M23 5' sequence,
mRNA sequence.
ACCESSION BG598154
VERSION BG598154.1 GI:13616294
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 659)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.
FEATURES
Location/Qualifiers
source
1..659
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cST520M23"
/clone_lib="cSTS"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
BASE COUNT 178 a 163 c 112 g 206 t
ORIGIN
Query Match 7.1%; Score 114.6; DB 155; Length 659;
Best Local Similarity 55.0%; Pred. No. 3.2e-18;
Matches 312; Conservative 0; Mismatches 244; Indels 11; Gaps 4:
QY 61 AGTCTTGAGAAATGCGGTGGCGCACCCAGCGCGTCG---CCGAGTTTACAGT 117
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 ACTGATTGACAAATGCAAGTTGCGGCACCTCCCGCGCGGCAACAGACTACGCCT 144
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 118 CCCACTGTGCTTTTCGACATGCGATGTTGATCTCTGATGACAGAACCATCTGCAATT 177
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 145 CCCTCTAACTTATTGTATCTATATGTTTCGCTTTTGGCTATATGCGTCGGATTATTTC 204
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 178 CTACAGATTCCGCCATCCCTTTGTCGAATC---TAAATTTATCATTTTCATCCAAATCGG 235
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 205 TACAAGTACCAATTTTCAAACTCGATTTTCGTTCAAAACCATTTATTCCTACTCTAAACAT 264
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 236 TCCCTTTCCCTGTCTCAACACATTTCTTCGTTAGCGGGGAATTTGATTTGGCGGTA 295
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 265 TCACCTCCCTCACTCTCAACACTACACACCTTAGCTGGCAACATGCTTGTGCCACTA 324
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 296 GATTCCTCCGATAGAAATGCGGAGTTGCGTTACAAAGAAAGGAGTCCCTTTCTTTAAACA 355
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 325 AAT---TCGAGTGGTTATCTCTGAGTGGCTTATGTGACTGGAGATTCATATATCTGTACT 381
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 356 AITGCAGATTCAGCATGGATTTTGATATCTCGCGGAGATCATCAGAGGATTCCTTAT 415
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 382 TTTATTGAGACTGATATGAATTTCAATCATCTCATTTGGTAACCATCTCGAAATGCTAAG 441
||| | ||| | ||| ||||| ||| ||||| ||| ||||| ||| |||||
QY 416 AAATTCACAGATTTGATTCGGCAGCTGCCAGACCGATTGTACCTCCGCGCAGCAAGTA 475
||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 GATTTTTATCCCTTTCATCTCTCAATTTGGCACACCTTAAAGTATGACCCGGGGGTCAAACTA 501
||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 476 TTACCACATTTTTTTCCTTACAGGTGACGGTGTCTCTCAACACCGGTATATGATTCATTTGACGC 535
||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 GTCCGGTCTTAGCCATTCAAGTGACACTTTTCCGAATCTTGGCATATCCCGTTGGTGT 561
||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 536 AATCT---TCATCAAGTCTTGTGTGATGCCAGTCTTCTTCTGCAATTTTAATAATATGG 592
||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 562 TTAGTAACCCCATCATGCTGCTTGTGATGAAATACCATCGTGAATTCATAAGAACATGG 621
||| ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 593 GTTTTGGTTGACAAATCCAAATGGAGAT 619
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 622 GGTTTACTCAACAAATTCGCGGTGAT 648
||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
BG097054
LOCUS 508 bp mRNA EST 29-JAN-2001
DEFINITION EST461573 potato leaves and petioles Solanum tuberosum cDNA clone
cSTB45P9 5' sequence, mRNA sequence.
ACCESSION BG097054
VERSION BG097054.1 GI:12587089
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 508)
van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
Generations of ESTs from potato leaves and petioles
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
Location/Qualifiers
source
1..508
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB45P9"
/tissue_type="potato leaves and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 127 a 130 c 87 g 164 t
ORIGIN
Query Match 6.9%; Score 112.4; DB 173; Length 508;
Best Local Similarity 55.5%; Pred. No. 1.1e-17;
Matches 281; Conservative 0; Mismatches 216; Indels 9; Gaps 3:
QY 68 GAGAAATGCGGTGTTGCGCCACCACCGGCGCGTCG---CCGAGTTTACAGTCCCACTG 124
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 GAGCAATGCCAAGTTGGCGCACCTCCCGACGCGCGGCAACAGACTACGCTCCCTCTA 61
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 125 TCGTTTTTCGCAATCGGATGGTTGATCTCTGATGACAGAACACCATCTGCAATTCACAGA 184
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 62 ACTTATTTTGATCATATATGCTCGGTTTGGCTATATGCTCGGATTTTATCTACAAG 121
Qy 185 TTCGCCATCTCTGCCAA---CTCTAAATTTATCATTTTCATCCATTAATTCGTCCTT 241
Db 122 CTACCAATTTCCAACTCGATTCGTTCTCAAAACCATTAATCTCTACTCTTAAACATTCACCTC 181
Qy 242 TCCCTTGTTCTCAACACATCTTCTTCGTTAGCCGGGAATTTGATTTGGCCGGTAGATTC 301
Db 182 TCCCTCACTCTCAACACACTACACACCTTAGCTGCAACATTTGCTTGCCACTAAAT--- 238
Qy 302 TCCGATAGAAATGCGGAGTTCGTTTACAGAAAGGGGACTCCGTTTCTTTTAAACAATTCGA 361
Db 239 TCGAGTGGTTATCTCTGATGCGTTATGTGACTGGAGATTCATATATCTGTTACTTTATT 298
Qy 362 GAATCGAGCATGGATTTGATTAATCTCCCGGAGATCATGAGGGATCTTATATAATTC 421
Db 299 GAGACTGATATGAATTTCAATCATCTCATTTGTTAACCATCTCGAAATGCTTAAGGATTTT 358
Qy 422 AACGATTTGATTCGCGAGCTGCCAGAACCGATTTGAACTTCGCGGACGAAAGTATTACCA 481
Db 359 TATCCCTTCATCTCAATTTGGCACAACCTAAGGATGACCCGGGGTCAAACTAGTCCCG 418
Qy 482 CTTTTTGGTTTACAGGTGACCGGTGTTCTTCAACACCGGTATATGATGAGGACGAATCTT 541
Db 419 GTCTTAGCCATTCAGTGACACTTTTTCGGAATCTTGGCATATCCGTTGGTTTAGTAAC 478
Qy 542 CATCAAGTCTTGTGTGATGCCAGTTC 567
Db 479 CATCATGTCGTTGTGATGGAATAC 504

RESULT 14
BE922746 574 bp mRNA EST 02-OCT-2000
LOCUS EST426515 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION CSTB21L12 5' sequence, mRNA sequence.
ACCESSION BE922746
VERSION BE922746.1 GI:10448822
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 574)
van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
Generation of ESTs from potato leaves and petioles
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
Location/Qualifiers
1. 574
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB21L12"
/tissue_type="petioles and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
162 a 123 c 106 g 183 t

BASE COUNT
ORIGIN
```

```
Query Match 6.7% Score 108.8; DB 142; Length 574;
Best Local Similarity 54.9% Pred. No. 9e-17;
Matches 281; Conservative 0; Mismatches 222; Indels 9; Gaps 3;

Qy 208 TAAATTTATCATTTTCATCCATTAATTCGTCCTCTTCCCTTCTCTCAACACACTTTCTTTCC 267
Db 63 TCAAAACCATTTATCTCTACTCTTAACATTTCACTCCCTCACTCTCAACACACTACACACC 122
Qy 268 GTTAGCCGGGAATTTGATTTGCGCGGTAGATTTCTCCCATAGAAATGCGGGAGTTGCGTTA 327
Db 123 CTTAGCTGGCAACATGCTTCTCCACTAAATTCGAGTGGTTAT---CCTGAGTTGCGTTA 179
Qy 328 CAAGAAGGGGACATCCGTTTCTTTAAACAATTCGAGAATTCGAGCATGGATTTTGATTAATCT 387
Db 180 TGTGACTGGAGATTCATATATCTGTTACTTTTATGAGACTGATATGAATTTCAATCATCT 239
Qy 388 CGCGGAGATCATCAGAGGGGATTTCTTATAAATTCACCAATTTGATTCGCGAGCTGCCAGA 447
Db 240 CATTTGGTAACCATCTCTCGAAATGCTAAGGATTT---TTATCCCATTTCTCAATTTGGCACA 296
Qy 448 ACCGATTTAACCTTCGCGGACGAGATTTACCACCTTTTTCGTTTACAGGTGACGGTGT 507
Db 297 ACCTAAGGATGACCGGGGGTCAAACTAGTCCGGTCTTAGCCATTTCAAGTGACACTTTT 356
Qy 508 CTCCAACACCCGTTATATGCATTTGGACGCAATCTTCACTCAAGTTCTTGTGTGATGCCAGTTC 567
Db 357 TCCGAATCTTGGCATATCCGTTGGTTTATGTAACCATCATGCTGCTGTGATGGAATAC 416
Qy 568 TTTTCTGCATTTTAAATAATTTGCGTTTGGTTTGACAAATCCAATCCAGATTCATTAAA 627
Db 417 CATCGTGAATTTCAAGAACAATGCGGTTTACTCAACAATAATTCGCGGTGAT---GAACA 473
Qy 628 GTTCTTCCACTTTCTCTCTACCTATCTAGACAGATCTCTGTTGCGAAGATCCATTTCA 687
Db 474 GTGCTTAGAGAAATGAGTTTCATTTCCATTTTATGATAGGTCCGTGATAAAGACCCCTTATGA 533
Qy 688 TATTCGTCGAAAAATCTACAATGAAAGAAAC 719
Db 534 ACAAGGGACGATTTATGCGGATGAAATGAAGC 565

RESULT 15
BF053202 403 bp mRNA EST 16-OCT-2000
LOCUS EST438432 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION CSTB34D1 5' sequence, mRNA sequence.
ACCESSION BF053202
VERSION BF053202.1 GI:10807098
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 403)
van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
Generation of ESTs from potato leaves and petioles
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
Location/Qualifiers
1. 403
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB34D1"
/tissue_type="leaflets and petioles"

FEATURES
source
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB34D1"
/tissue_type="leaflets and petioles"
```

```
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/notor="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT      110 a      81 c      78 g      134 t
ORIGIN

Query Match      6.4%; Score 103.6; DB 143; Length 403;
Best Local Similarity 57.6%; Pred. No. 1.7e-15;
Matches 228; Conservative 0; Mismatches 159; Indels 9; Gaps 2;

QY 209 AAATTTATCATTTTCATCCCAATTAATCGTCCTTTCCCTTGTCTCTCAACACTTTCTTCGG 268
   ||| || ||| | | |||| | | |||| | | |||| | | |||| | |
Db 11 AAACCAATTAATTCCTACTCTTAACACTTCACTCTCCCTCGCTCTCCAACACTATATGATT 70
   ||| || ||| | | |||| | | |||| | | |||| | | |||| | |
QY 269 TTAGCCGGGAATTTGATTTGGCCGGTAGATTCTCGGATAGAAATGCGGAGTTGCGTTAC 328
   |||| | | ||| | | ||| | | ||| | | ||| | | ||| | | ||| | |
Db 71 TTAGCCGGTAATGTGTTTATCGTTAAATTCAAACAACCTTA---CCTGAATTCGGTTAT 127
   |||| | | ||| | | ||| | | ||| | | ||| | | ||| | | ||| | |
QY 329 AAGAAAGGGGACTCCGTTTCTTTAACAATTCGAGAATCGAGCATGGATTTGATTATCTC 388
   |||| | | ||| | | ||| | | ||| | | ||| | | ||| | | ||| | |
Db 128 GTAACAGGGGATTCTGTGCAGTTACTTTTCTGAAAGTGATGGATTTCAATTTTCTC 187
   |||| | | ||| | | ||| | | ||| | | ||| | | ||| | | ||| | |
QY 389 GCCGGAGATCATCAGAGGGATTCTTATAATTCACGATTTGATTCGGCAGCTGCCAGAA 448
   |||| | | ||| | | ||| | | ||| | | ||| | | ||| | | ||| | |
Db 188 ATAGGAGACCATCCTAGAAACGCAAGGATTTTATACATTTGTTCCTCAAGTTGGCGGAA 247
   |||| | | ||| | | ||| | | ||| | | ||| | | ||| | | ||| | |
QY 449 CC-----GATTGTAACCTCCGGCGACGAAGTATTACCACTTTTGTGTTTACAGGTGACG 502
   || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 248 CCTTGTAAGGCAAGCGGGGATGTCCATATTAGCCCCCTTGTAGCCATTCAAGTGACA 307
   || | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 503 GTGTTCTCCAACACCGGTATATGATTTGGACGCAATCTTCATCAAGTCTGTGTGATGCC 562
   || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 308 CTTTTTCCGAATCATGCGATATCGATTGGTTTCACTAACCATCATGTTGTTGGTGACGGA 367
   || | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 563 AGTCTCTTTCTGCATTTTATAAATTAATATGGTTTG 598
   || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 368 CTAGCATAGAGGGTTTATAAAGCATGGGCTATG 403
   || | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: November 5, 2001, 18:00:43
Job time: 14883 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:03:41 ; Search time 168.74 seconds
(without alignments)
1819.738 Million cell updates/sec

Title: US-08-894-356C-2
Perfect score: 1622
Sequence: 1 GAACCATGTAATCCAAATTA.....AAAAAAAAAAAAAAAAAAAA 1622

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/lna/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40.8	2.5	7218	1	US-08-232-463-14
2	39.6	2.4	51952	3	US-08-947-823-1
3	38	2.3	1953	1	US-08-463-975-1
4	38	2.3	1953	1	US-08-783-266-1
5	38	2.3	1953	2	US-09-023-327-1
6	37.6	2.3	19124	2	US-08-487-826B-13
7	37.6	2.3	636	4	US-08-998-416-1137
8	37.6	2.3	837	4	US-08-998-416-288
9	37.4	2.3	615	4	US-08-998-416-186
10	37	2.3	575	1	US-08-554-659-11
11	37	2.3	2821	4	US-09-103-429A-2
12	36.4	2.2	4160	4	US-09-341-678-5
13	35.8	2.2	894	3	US-08-906-769-90
14	35.8	2.2	894	3	US-08-906-616-90
15	35.8	2.2	894	3	US-08-817-795-90
16	35.8	2.2	894	3	US-08-639-075A-90
17	35.8	2.2	894	4	US-09-012-431-90
18	35.8	2.2	894	4	US-09-012-692-90
19	35.8	2.2	894	4	US-08-906-613-90
20	35.8	2.2	894	5	PCT-US95-14442A-90
21	35.6	2.2	602	1	US-08-764-100-8
22	35.6	2.2	642	1	US-08-764-100-13
23	35.6	2.2	643	1	US-08-764-100-7
24	35.6	2.2	2993	1	US-08-764-100-2
25	35.6	2.2	2993	1	US-08-764-100-10
26	35.6	2.2	3000	1	US-08-764-100-9
27	35.6	2.2	3001	1	US-08-764-100-1

Sequence 16, Appl
Sequence 23, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 67, Appl
Patent No. 5194596
Patent No. 5219739
Sequence 1, Appl
Sequence 32, Appl
Sequence 32, Appl
Sequence 32, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 9, Appl
Sequence 9, Appl
Sequence 1, Appl
Sequence 1, Appl

US-09-128-155-16
US-08-117-083-23
US-08-207-412B-1
PCT-US95-02950-1
US-08-117-083-67
5194596-16
5219739-16
US-09-286-690-1
US-07-991-867B-32
US-08-107-755A-32
US-08-544-332-32
US-07-991-867B-8
US-08-107-755A-8
US-08-544-332-8
US-08-021-608D-9
US-08-726-160-9
PCT-US94-01782-9
US-08-021-608D-1

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EF 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14


```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-487-826B-13

Query Match      2.3%; Score 38; DB 2; Length 19124;
Best Local Similarity 53.3%; Pred. No. 0.97;
Matches 80; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 1473 AAGAAAAGTGGTATCAATGTATATAAAGACAGACAGCAAGTTATGATGCAACAAATGTTT 1532
Db 15686 AAAATATATATCATATAAACAACAAAAAAGAAAAAATATATATAAATAAATATATA 15745
QY 1533 TAGGAGATTACAAATCCATGGAAGATGTATCAAACTCATCTCTATATATATATATTC 1592
Db 15746 TCATAAATAAATAAATAATTAATAAATGTTAAAAAATAATATACATAAATAAATAA 15805
QY 1593 AATTGTTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1622
Db 15806 AAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15835

RESULT 7
US-08-998-416-1137/c
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1692RP
; US-08-998-416-1137

Query Match      2.3%; Score 37.6; DB 4; Length 636;
Best Local Similarity 55.3%; Pred. No. 0.21;
Matches 73; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1490 AATGTATATAAAGACAGACAGCAAGTTATGATGCAACAAATGTTTATAGGAGATTACAAATCC 1549
Db 443 ATTAATAATATAAATCAACATAATATTTATAAATAAGATATATATAAATAAATAAATAA 384
QY 1550 ATGGGAAGATGTATCAAACTCATCTCTATATATATATATATATATATCAATTTGTTTTAAAAA 1609
Db 383 TTTACAATATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 324
QY 1610 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1621
Db 323 ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 312

RESULT 8
US-08-998-416-288/c
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebischung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
```


[illegible]


```

; APPLICANT: Rushlow, Keith E.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Frank, Glenn R.
; APPLICANT: Heath, Andrew W.
; APPLICANT: Yamaka, Miles Yamanaka
; APPLICANT: Arfsten, Ann
; APPLICANT: Dale, Beverly
; APPLICANT: Stiegler, Gary
; TITLE OF INVENTION: USE OF PROTEASE INHIBITORS AND
; TITLE OF INVENTION: PROTEASE VACCINES TO PROTECT ANIMALS FROM FLEA
; TITLE OF INVENTION: INFESTATION, AND FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, AND USES THEREOF
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,795
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/14442
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gary J. Connell
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..766
; FEATURE:
; NAME/KEY: Xaa - any amino acid
; LOCATION: 187
; US-08-817-795-90

Query Match 2.2%; Score 35.8; DB 3; Length 894;
Best Local Similarity 49.7%; Pred. No. 0.81;
Matches 91; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1439 TTCATCAATAGTCTCCATGTGAGGAGCCCTTTGTAAGAAAAAGTGGTATCAATGTATAA 1498
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 709 TTACCAATGTTGCTCATTTTGTGGATTGGAATCAAGAACATAGTGGAAATTGTAA 768

QY 1499 AAAAGACAGACAAGTATGATGCAACAATGTTTATAGGAGATTACAAATCCATGGGAAGA 1558
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 769 AATAAACTCAAACTAAAAAAAATAATTAATGCACTGAAAAATTTTCAAGAAAA 828

QY 1559 TGTATCAAACTCATCTCTATATATATATATCAATGTTTAAAAAATAAAAAA 1618
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 829 GTTTGGATCGTTTGTAAATGTAATGACATAAAAGCGTAAATTAGAAAAA 888

QY 1619 AAA 1621
|||

```

Db 889 AAA 891

Search completed: November 5, 2001, 18:04:15
Job time: 15055 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:11:15 ; Search time 445.49 seconds
(without alignments)
2286.149 Million cell updates/sec

Title: US-08-894-356c-2

Perfect score: 1622

Sequence: 1 GAACCAATGAATCAATTA.....AAAAAAAAAAAAAAAAAAAA 1622

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601.*
1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1607.6	99.1	1622	17 AAT37309
2	218	13.4	1479	17 AAT37311
3	190.8	11.8	1703	17 AAT37308
4	153.8	9.5	1518	17 AAT37313
5	139.8	8.6	1508	17 AAT37312
6	96.8	6.0	936	22 AAF58252
7	96.8	6.0	936	22 AAF58254
8	96.8	6.0	936	22 AAF58257
9	96.8	6.0	936	22 AAF58259
10	96.8	6.0	936	22 AAF58262
11	96.8	6.0	938	22 AAF58255

12	95.4	5.9	936	22	AAF58252	Oligonucleotide D1
13	95.4	5.9	935	22	AAF58254	Oligonucleotide D1
14	95.4	5.9	935	22	AAF58257	Oligonucleotide D1
15	95.4	5.9	935	22	AAF58259	Oligonucleotide D2
16	95.4	5.9	936	22	AAF58262	Oligonucleotide D2
17	95.4	5.9	938	22	AAF58255	Oligonucleotide D1
18	58.4	3.6	244	22	AAF58238	Oligonucleotide D1
19	56.2	3.5	244	22	AAF58238	Oligonucleotide D1
20	47.8	2.9	2969	21	AAA23450	cDNA encoding huma
21	47.2	2.9	1240	21	AAC53227	Arabidopsis thalia
22	45.8	2.8	328	21	AAA31809	Plant microsatelli
23	43.8	2.7	280	16	AAT21000	Human gene signatu
24	43.8	2.7	1273	21	AAC80571	Human secreted pro
25	43.4	2.7	1595	10	AAN90733	Genomic DNA from P
26	43.4	2.7	5376	10	AAN90732	Sequence of plasm
27	42.4	2.6	7458	21	AA70106	Plasmodium falcipa
28	42	2.6	1982	10	AAN90225	Malaria-specific P
29	41.4	2.6	471	21	AAA16053	Human colon cancer
30	41.4	2.6	1523	21	AAC59768	Human secreted pro
31	41.4	2.6	1823	21	AAC78037	Human cancer assoc
32	41.4	2.6	2851	20	AAC33599	Human breast tumou
33	41.2	2.5	3975	9	AAN81157	Malaria-specific g
34	41.2	2.5	3975	13	AAQ22999	SERP gene. Plasm
35	41.2	2.5	6124	11	AAQ03568	Sequence encoding
36	40.8	2.5	9789	17	AAT41852	cDNA encoding Plas
37	40.4	2.5	769	18	AAT79095	Full length human
38	40	2.5	13499	21	AAF22286	BAC containing rep
39	39.8	2.5	498	22	AAF68733	Human lung tumour
40	39.8	2.5	969	21	AAC77818	Human cancer assoc
41	39.8	2.5	3366	22	AAF24684	Nucleotide sequenc
42	39.8	2.5	3366	22	AAF24706	Nucleotide sequenc
43	39.8	2.5	3515	22	AAF5343	Human TGF-beta rec
44	39.8	2.5	10442	22	AAF24680	Nucleotide sequenc
45	39.8	2.5	10442	22	AAF24702	Nucleotide sequenc

ALIGNMENTS

RESULT 1

AAT37309
ID AAT37309 standard; cDNA to mRNA; 1622 BP.

XX AC AAT37309;

XX DT 06-FEB-1997 (first entry)

XX DE Aromatic acyl transferase coding sequence.

XX KW Aromatic acyl transferase; transformation; anthocyanin pigment;

XX KW plants; acylation; colour; tone; colouration; colour change;

XX KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;

XX KW Scenecio cruentus; Lavandula angustifolia; ds.

XX OS Gentiana triflora var. Japonica (Clone pGAT106).

XX FH Key Location/Qualifiers

XX CDS 6..1415

XX FT /*tag= a

XX FT /product= Aromatic acyl transferase.

XX FT 701..703

XX FT /*tag= b

XX FT /transl_except= ATC encodes Leu

XX FT 1322..1324

XX FT /*tag= c

XX FT /transl_except= AAT encodes Asp

XX PN WO9625500-A1.

XX XX 22-AUG-1996.

XX PF 16-FEB-1996; 96WO-JP00348.

XX XX

Db 1027 gcyttgcaaaagcaacacataaaagagtagttggggataaaaggcttcttggtgcagttg 1086
 QY 1119 CGCCATCGGAGATGCTATCCATAGAGGTTACATGACTACGAGGATTTGAGAGGAG 1178
 Db 1087 cagctattggagaagccattgaaaagaggttgcaacaacaaaagcgttcttgagatg 1146
 QY 1179 APTGATATCGCGCCCGGATCAACATCTGCGGCACCAAGGTCGACGCTCATTTATGTCG 1238
 Db 1147 caaaaacttggttaccgaatcaat-----ggaatccctcaaaaagattctcg 1197
 QY 1239 TTGGATCGGCACACGCAATGTCAGATGTTGATGTCAGATTTTGGTGGGAAAGCTTG 1298
 Db 1198 ggaattaccggatcgctaaagtctgattcgatggtgtagatttgatggggaaagcctg 1257
 QY 1299 AAACCATGATCTCTTCAACTAATCTTCGGCAACACTAATTTTGATCTCTCGGTCCA 1358
 Db 1258 caaatttgaca---ttacctctgttgattatgcagaattgattatgattcagtcaca 1314
 QY 1359 GAAGATTTAAAGGAGCAGCTTGAGCTTGGCATTTCTTGGCTAAGAAATAGGATGGACGAT 1418
 Db 1315 gggatttgaaaagagtgagatggagtagatcattgcctaagattcattatgagatgat 1374
 QY 1419 TTGCCACCATTTT 1432
 Db 1375 ttgcaaaaatcttt 1388

RESULT 4

AAT37313
 ID AAT37313 standard; cDNA to mRNA; 1518 BP.
 XX
 AC AAT37313;
 XX
 DT 06-FEB-1997 (first entry)
 XX
 DE Aromatic acyl transferase coding sequence.
 KW Aromatic acyl transferase; transformation; anthocyanin pigment;
 KW plants; acylation; colour; tone; colouration; colour change;
 KW Gentiana triflora; Petunia hybrida; Perilla octinoides;
 KW Scenecio cruentus; Lavandula angustifolia; ds.
 XX
 OS Lavandula angustifolia (Clone pLAT21).
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1355
 FT /*tag= a
 FT /product= Aromatic acyl transferase.
 XX
 PN W09625500-A1.
 XX
 PD 22-AUG-1996.
 XX
 PF 16-FEB-1996; 96WO-JP00348.
 XX
 PR 30-JAN-1996; 96JP-0046534.
 PR 17-FEB-1995; 95JP-0067159.
 PR 29-JUN-1995; 95JP-0196915.
 XX
 PA (SUNR) SUNTORY LTD.
 XX
 PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
 PI Nakao M, Tanaka Y, Yonekura K;
 XX
 DR WPI; 1996-393401/39.
 XX
 PT DNA coding for aromatic acyl transferase - for transforming plants
 PT which produce anthocyanin pigments and thus altering colour tone,
 PT e.g. of flowers
 XX
 PS Claim 4; Page 73-76; 94pp; Japanese.
 XX

CC Vectors containing DNA fragments encoding proteins of plant origin
 CC with aromatic acyl transferase activity may be used to transform
 CC plants which produce anthocyanin pigments. The aromatic acyl
 CC transferase acylates the pigments in the flower resulting in colour
 CC tone changes and allowing new colourations to be produced. Six
 CC specific DNA sequences encoding aromatic acyl transferase from
 CC different plants are described in AAT37308-T37313. NOTE: This
 CC sequence is supposed to cross reference with the protein described
 CC in AAW04727, however there are so many discrepancies between the
 CC polypeptide decoded from this sequence and the polypeptide given in
 CC the specification and described in AAW04727 that the indexer decided
 CC not to cross reference the two.
 XX

SQ Sequence 1518 BP; 384 A; 340 C; 366 G; 428 T; 0 other;

Query Match 9.5%; Score 153.8; DB 17; Length 1518;
 Best Local Similarity 49.5%; Pred. No. 8.4e-30;
 Matches 686; Conservative 0; Mismatches 642; Indels 57; Gaps 9;

QY 72 AATGCGGTGTTGCGCCACCACCGGACGCGTCCGAGTTTACAGTCCACATGTCGTTT 131
 Db 16 aatcccgagtgccgcgcgcctccagggcagcgtggtgagcagtcactccgcaccctct 75
 QY 132 TCGACATCGATGTTGATCTCTGATGCAGAACACATCTGCAATTTCTACAGATTCGCC 191
 Db 76 tcgacatgacgtggtgcacattccaccccatgcttcagcttctcttcagaaacctct 135
 QY 192 ATCTGTGTCGCAACTCTA---AATTTATCATTTTCATCCATTAATCGTCCCTTTCCCT 248
 Db 136 gtccaaaacccgcttctcctgaaacgctgttccgaaacactcaacaactctctctaa 195
 QY 249 TTCTCAACACATTTCTCCGTTAGCGGGAATTTGATTTGGCGGTAGATTCCTCGGATA 308
 Db 196 cctcaaacactctctccctcttcacatcactcactcactcactcactcactcactcact 252
 QY 309 GAATGCGGAGTTGGTTACAGAAAGGGGACTCCGCTTTCTTTAAACAATTCAGCAATCGA 368
 Db 253 aaatgccgagttccggtatcagaaacggtgactcgttctcttcacgattatggagctg 312
 QY 369 GCATGGATTTGATTTATCTCGCGGAGATCATCAGAGGATTTCTTATAAATTCACAGATT 428
 Db 313 tcggagatcctccgctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 372
 QY 429 TGATTCGCGAGCTGCAGAACCGATTGTAACCTCGCGGACGAAGTATTACACATTTTGG 488
 Db 373 aagatctccagctgcgcgcgtagtcgaggaatctgtagcgaatgtttcagattttag 432
 QY 489 CTTTACAGTGACGGTGTCTCCAAACACCGGTATATGCAATTCGAGCAATCTTCATCAAG 548
 Db 433 ccgtgcaagtactctgtttcccggtcgcggtggtgcatcggaataacgacgcaccaca 492
 QY 549 TTCTTGGTGATGCCAGTTCTTTCTGCAATTTTAAATAATTTATGTTGGTGGTGGTGA 608
 Db 493 ccgttagcagatgctccatcgttttaggtttatgaagagttgggtccctcactcaaat 552
 QY 609 CCAATGGAGATTC---ATTAAAGTTCCCTCCACTTCTTCTCTACCTATGACGACAGAT 665
 Db 553 tcggagagatgatgaattcttggacggaaaagggtgaatgtttgcgggttttcgacgat 612
 QY 666 CTGTGTGCAAGATCCATTTTCATATTCGTGAAATAATCTACAAATGAAGAAACTCTCTCA 725
 Db 613 cgtcgtgaattatccgccthaattggacacataattatggaaac-----aacgcgcaga 666
 QY 726 AATCTCAGGGCACACTACTGTTCTAAATTCAGCAATTTCTAAAGATGAAGTTTCGAGCCA 785
 Db 667 aacgtccgttggaaac-----gcagcatccatctttaccgcgagtcggttcgagcta 720
 QY 786 CTTTCATCTACACCTATTGATATCATCAAGCTCAAGAAATTCATTTCTGTCAAAAATC 845
 Db 721 cctactctttcccccactcgaattgaagaaattgaaaggtttgattcagagaaagccc 780
 QY 846 GCAACTTAACCGGTAGTAGTAATATAATCTGTCAACTTTTACGGGTGACATCTGCACCTGA 905

[illegible]

RESULT 6
AAF58252/C
ID AAF58252 standard; DNA; 936 BP.

XX AAF58252;

XX
DT 24-APR-2001 (first entry)XX
DE oligonucleotide D1835.

XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.

XX OS Synthetic.

XX PN WO200107665-A2.

XX

PD	01-FEB-2001.
XX	
XX	26- JUL- 2000; 2000WO-US20476.
PF	
XX	
XX	26- JUL- 1999; 99US-0145695.
PR	
XX	17-MAR-2000; 2000US-0190259.
PR	
XX	
PA	(CLIN-) CLINICAL MICRO SENSORS INC.
XX	
XX	Umek RM;
PI	
XX	
XX	WPI; 2001-159728/16.
DR	
XX	
XX	Nucleic acids containing electron-transfer group, useful as labels in
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT	a single surface -
XX	
XX	
PS	Example 6; Page 127; 159pp; English.
XX	
CC	The present invention relates to a composition comprising two nucleic
CC	acids each containing an electron-transfer group (ETW) having
CC	different redox potentials. The invention is used for electronic
CC	detection of nucleic acids, especially of substitutions (mismatches)
CC	and single-nucleotide polymorphisms, e.g. for genotyping,
CC	monitoring gene expression.
XX	
XX	
SQ	Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match 6.0%; Score 96.8; DB 22; Length 936;
Best Local Similarity 1.0%; Pred. No. 2.8e-15;
Matches 8; Conservative 453; Mismatches 305; Indels 0; Gaps

[illegible]

```
QY 1457 GTGAGGAGCCCTTTGTGAAGAAAAAGTGGTATCAATGTATAAAAAAGACACACAAGTTAT 1516
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 131
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1517 GATCAACAATGTTTGTAGGAGATTACAATCCATGGAAGATGATCAACATCATCTCT 1576
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 71
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1577 CTATATATATATATCAATGTTTAAAAAAAATAAAAAAATAAAAAA 1622
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 25
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 7
AAF58254/c
ID AAF58254 standard; DNA; 936 BP.
XX
AC AAF58254;
DT 24-APR-2001 (first entry)
DE Oligonucleotide D1875.
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
XX WO200107665-A2.
XX
PD 01-FEB-2001.
XX
XX 26-JUL-2000; 2000WO-US20476.
XX
XX 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
XX
XX WPI; 2001-159728/16.
XX

Nucleic acids containing electron-transfer group, useful as labels in
hybridization assays, e.g. for genotyping, allowing repeat analyses on
a single surface
Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 6.08; Score 96.8; DB 22; Length 936;
Best Local Similarity 1.08; Pred. No. 2.8e-15;
Matches 8; Conservative 453; Mismatches 305; Indels 0; Gaps 0;

QY 857 GGTAGTAGTATTATCTGTCAACTCTTCACGGTGACATCTGCACATGCTGGACATGC 916
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 790 GGCWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 731
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 917 TTGTCGAAATCATTAGACACCGTCGTAAGAGAGAGAGTGAAGAGGATAAACATGCAGCA 976
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 730 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 671
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 977 AACTTATGCTGCTTCATCACTGCCGACACAGGTTTCTGCTCGCGGATACCTCAAAATAC 1036
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
Db 670 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 611
QY 1037 TTTGGAATTCGATAGTGCCTTGTATGTTGGGATCGACTCATGAGCAACTTGTAGGAAAT 1096
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 610 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 551
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1097 GAAGGGTTGCGGTAGTGCACACGCCCATCGGAGATGCTATCCATAAGAGGTTACATGAC 1156
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 550 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 491
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1157 TACGAAGAAATCTGAGAGGAGATGCGCCGCCGATCATCAACATCTGCGGCACCA 1216
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 490 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 431
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1217 AGTTCGACGCTCATTTATGTCGTTGGATCCGACACGCAATGTGCATGATTTTGTATGCA 1276
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 430 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 371
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1277 GATTTTGGTTGGGAAAGCTTGAAAACATGAATCTGTTTCAACTAATCTCTCGGCAACA 1336
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 311
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1337 CTAATTTTGATCTCTCGTCCGACGAAGATTTAAAGGACCACTTGAGCTTGGCATTTCTTG 1396
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 310 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 251
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1397 CCTAAGAAATAGGATGACGCACTTTCACCAATTTTACGAATTTTCATCAATAGTCTCCAT 1456
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 250 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 191
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1457 GTGAGGAGCCCTTTGTGAAGAAAAAGTGGTATCAATGTATAAAAAAGACACACAAGTTAT 1516
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 190 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 131
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1517 GATCAACAATGTTTGTAGGAGATTACAATCCATGGAAGATGATCAACATCATCTCT 1576
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 130 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 71
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1577 CTATATATATATATCAATGTTTAAAAAAAATAAAAAAATAAAAAA 1622
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 70 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 25
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 8
AAF58257/c
ID AAF58257 standard; DNA; 936 BP.
XX
AC AAF58257;
XX
XX 24-APR-2001 (first entry)
DE Oligonucleotide D1954.
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
XX Synthetic.
XX
XX WO200107665-A2.
XX
PD 01-FEB-2001.
XX
XX 26-JUL-2000; 2000WO-US20476.
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
XX
XX WPI; 2001-159728/16.
XX
```


[illegible]

Db 84 www..... 143

Qy 248 GTTCTCAACACTTCTTCGGTAGCCGGGAATTGATTTGGCGGTAGATCTCCTCCGAT 307

Db 144 www..... 203

Qy 308 AGAATGCCGGAGTTGCGTTACAAGAGGGGACTCCGTTTCTTAACAATTTGCAGAAATCG 367

Db 204 www..... 263

Qy 368 AGCATGGATTTGATTTATCTCGCGGAGATCATCAGAGGGGATTTCTTATAAATTTCAACGAT 427

Db 264 www..... 323

Qy 428 TTGATTCGCGAGTGCACAGACCGATTTGTAACCTCCGCGCAGAGATTTACCACTTTT 487

Db 324 www..... 383

Qy 488 GCTTTACAGTGACGGTGTCTCCAACACCGGTATATGATTTGGACGCAATCTTTCATCAA 547

Db 384 www..... 443

Qy 548 GTTCTGTGTATGATCGGTTCTTCTGCAATTTTAATAAATTTATGGGTTTGGTTGACAAA 607

Db 444 www..... 503

Qy 608 TCCATGGAGATTCATTAAGCTTCTCCACTTCTCTACCTATGCTACGACAGATCT 667

Db 504 www..... 563

Qy 668 GTGTGCAAGATCCATTTTCATATTGCTGGAATAATCTACAATGAAGAAACTGCTCAAA 727

Db 564 www..... 623

Qy 728 TCTCAGGCACACCTACTGTTCTTAATCCACCAATTTCTAAGATGAAGTTTCGAGCCACC 787

Db 624 www..... 683

Qy 788 TTCATCTACACCTATTTGATATCATGAAGCTCAAGAAATTTCAATTCGTCAAAAATCGC 847

Db 684 www..... 743

Qy 848 AACTTAACCGGTAGTAGTAATTAATCTGTCAACTTTTCACGGTGACATCTGCAC 902

RESULT 13

AAF58254

ID AAF58254 standard; DNA; 936 BP.

XX AAF58254;

XX 24-APR-2001 (first entry)

XX Oligonucleotide D1875.

XX Electron-transfer group; ETM; mismatch; genotyping;

XX gene expression; ss.

XX Synthetic.

XX WO200107665-A2.

XX 01-FEB-2001.

XX 26-JUL-2000; 2000WO-US20476.

XX 26-JUL-1999; 99US-0145695.

XX 17-MAR-2000; 2000US-0190259.

XX (CLIN-) CLINICAL MICRO SENSORS INC.

PI Umek RM;

XX WPI; 2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in

PT hybridization assays, e.g. for genotyping, allowing repeat analyses on

PT a single surface

XX Example 6; Page 127; 159pp; English.

XX The present invention relates to a composition comprising two nucleic

CC acids each containing an electron-transfer group (ETM) having

CC different redox potentials. The invention is used for electronic

CC detection of nucleic acids, especially of substitutions (mismatches)

CC and single-nucleotide polymorphisms, e.g. for genotyping,

CC monitoring gene expression.

XX Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 5.9%; Score 95.4; DB 22; Length 936;

Best Local Similarity 1.5%; pred.No. 6.3e-15;

Matches 12; Conservative 451; Mismatches 312; Indels 0; Gaps 0;

Qy 128 TTTTTCGACATCGATGGTTGATCTCTGATGCAGAACACCATCTGCATTTTACAGATTC 187

Db 24 www..... 83

Qy 188 CGCATCTCTTCCCACTCTAAATTTATCATTTTCATCCATTAATTCGCTTTCCTT 247

Db 84 www..... 143

Qy 248 GTTCTCAACACTTCTTCCGTTAGCCGGGAATTTGATTTGCGCGGTAGATTTCTCCGAT 307

Db 144 www..... 203

Qy 308 AGAATGCCGGAGTTGCGTTACAAGAGGGGACTCCGTTTCTTAACAATTTGCAGAAATCG 367

Db 204 www..... 263

Qy 368 AGCATGGATTTGATTTATCTCGCGGAGATCATCAGAGGGGATTTCTTATAAATTTCAACGAT 427

Db 264 www..... 323

Qy 428 TTGATTCGCGAGTGCACAGACCGATTTCTAACCCTCCGCGCAGAGATTTACCACTTTT 487

Db 324 www..... 383

Qy 488 GCTTTACAGTGACGGTGTCTCCAACACCGGTATATGATTTGGACGCAATCTTTCATCAA 547

Db 384 www..... 443

Qy 548 GTTCTGTGTATGCGGATCTTCTTCTGCAATTTTAATAAATTTATGGGTTTGGTTGACAAA 607

Db 444 www..... 503

Qy 608 TCCATGGAGATTCATTAAGCTTCTTCCACTTCTCTACCTATGCTACGACAGATCT 667

Db 504 www..... 563

Qy 668 GTGTGCAAGATCCATTTTCATATTGCTGGAATAATCTACAATGAAGAAACTGCTCAAA 727

Db 564 www..... 623

Qy 728 TCTCAGGCACACCTACTGTTCTTAATCCACCAATTTCTAAGATGAAGTTTCGAGCCACC 787

Db 624 www..... 683

Qy 788 TTCATCTACACCTATTTGATATCATGAAGCTCAAGAAATTTCAATTCGTCAAAAATCGC 847

Db 684 www..... 743

Qy 848 AACTTAACCGGTAGTAGTAATTAATCTGTCAACTTTTCACGGTGACATCTGCAC 902

CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

Query Match		5.9%;	Score 95.4;	DB 22;	Length 936;
Best Local Similarity		1.5%;	Pred. No. 6.3e-15;		
Matches 12;		Conservative 451;	Mismatches 312;	Indels 0;	Gaps 0;
Qy	128	TTTTTCGACATGCGATGTTGATCTCTGATGCAGAACACCATCTGCATTTCTACAGATTC	187		
Db	24	TTTTTCGACATGCGATGTTGATCTCTGATGCAGAACACCATCTGCATTTCTACAGATTC	83		
Qy	188	CGCATCCTTCTCCCAACTTAATTTATCATTTTCATCCATTAATCGTCCCTTCCCTT	247		
Db	84	CGCATCCTTCTCCCAACTTAATTTATCATTTTCATCCATTAATCGTCCCTTCCCTT	143		
Qy	248	GTCTCTAAACACACTTTCTTCCGTTAGCCGGGAATTTGATTTGGCCGGTAGATTCCTCCGAT	307		
Db	144	GTCTCTAAACACACTTTCTTCCGTTAGCCGGGAATTTGATTTGGCCGGTAGATTCCTCCGAT	203		
Qy	308	AGATGCCGGAGTTGCGTTACAAAGGAGGACTCCGTTCTTTAAACAATTCAGAAATCG	367		
Db	204	AGATGCCGGAGTTGCGTTACAAAGGAGGACTCCGTTCTTTAAACAATTCAGAAATCG	263		
Qy	368	AGCATGGATTTTGATTAATCTCGCCGGAGATCATCAGAGGGATTCTTATAAATTCACGAT	427		
Db	264	AGCATGGATTTTGATTAATCTCGCCGGAGATCATCAGAGGGATTCTTATAAATTCACGAT	323		
Qy	428	TGATTTCCGACAGTCCAGAACCGATTGTAACCTCCGGCGAGAGTATTACCACTTTT	487		
Db	324	TGATTTCCGACAGTCCAGAACCGATTGTAACCTCCGGCGAGAGTATTACCACTTTT	383		
Qy	488	GCTTACAGTGACGGGTTCTCCAAACCGGATATGATGCATGGACGCAATCTTCATCAA	547		
Db	384	GCTTACAGTGACGGGTTCTCCAAACCGGATATGATGCATGGACGCAATCTTCATCAA	443		
Qy	548	GTTCTTGATGTCAGTTCCTTTCTGATTTTAATAAATTTATGGTTTGGTTGACAAA	607		
Db	444	GTTCTTGATGTCAGTTCCTTTCTGATTTTAATAAATTTATGGTTTGGTTGACAAA	503		
Qy	608	TCCATGGAGATTCATTAAGTTCCTTCCACTTTCTTCTACCTATGTACGACAGATCT	667		
Db	504	TCCATGGAGATTCATTAAGTTCCTTCCACTTTCTTCTACCTATGTACGACAGATCT	563		
Qy	668	GTGGTGCAGATCCATTTTCATATTCGTCGAAAAATCTACAATGAAGAAACTGCTCAA	727		
Db	564	GTGGTGCAGATCCATTTTCATATTCGTCGAAAAATCTACAATGAAGAAACTGCTCAA	623		
Qy	728	TCTCAGGACACCTACTGTTCTAAATCCAGCAATTTCTAAAGATGAAGTTCGAGCCACC	787		
Db	624	TCTCAGGACACCTACTGTTCTAAATCCAGCAATTTCTAAAGATGAAGTTCGAGCCACC	683		
Qy	788	TTCATCTTACACCTATTTGATATCATGAAGCTCAAGAAATTCATTTCTCAAAAAATCGC	847		
Db	684	TTCATCTTACACCTATTTGATATCATGAAGCTCAAGAAATTCATTTCTCAAAAAATCGC	743		
Qy	848	AACTTAACCGGTAGTAGTAATTAATCTGTCACACTTTCACGGTGACATCTGCAC	902		
Db	744	AACTTAACCGGTAGTAGTAATTAATCTGTCACACTTTCACGGTGACATCTGCAC	798		

Search completed: November 5, 2001, 18:11:26
Job time: 15421 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 16:30:46 ; Search time 8904.87 Seconds
(without alignments)
2817.412 Million cell updates/sec

Title: US-08-894-356C-2
Perfect score: 1622
Sequence: 1 GAACCATTTGAATCCAATTAA.....AAAAAAAAAAAAAAAAAAAA 1622

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v1:*
59: gb_v2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_ro1:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1622	100.0	1622	10	E12754
2	1622	100.0	1622	12	AB026494
3	218	13.4	1479	10	E12756
4	216	13.3	1476	12	AB029340
5	189.2	11.7	1679	12	AB010708
6	189.2	11.7	1703	10	E12753
7	139.8	8.6	1508	10	E12757
8	50.4	3.1	178273	60	AC005308

c 9	50.2	3.1	205429	60	AC005506	AC005506 Plasmid
c 10	50	3.1	169794	60	AC004688	AC004688 Plasmid
c 11	50	3.1	196149	60	AC004709	AC004709 Plasmid
c 12	49	3.0	42781	14	MISOLSUR	X17375 Scenedesmus
c 13	49	3.0	42919	13	AF204057	AF204057 Scenedesmus
c 14	49	3.0	234112	96	PFMAL4P2	AL035475 Plasmid
c 15	48.2	3.0	144457	67	AC023268	AC023268 Homo sapi
c 16	48	3.0	101176	12	AC002560	AC002560 Genomic s
c 17	46.8	2.9	1337	88	AF090922	AF090922 Homo sapi
c 18	46.8	2.9	66237	12	AB016892	AB016892 Arabidops
c 19	46.6	2.9	85779	14	SC601856	AJ011856 Saccharom
c 20	46.6	2.9	178273	60	AC005308	AC005308 Plasmid
c 21	46.4	2.9	153477	60	AC006278	AC006278 Plasmid
c 22	46.4	2.9	192581	84	PFMAL13P1	AL049180 Plasmid
c 23	46.2	2.8	83110	84	PFMAL13PB	AL109814 Plasmid
c 24	46.2	2.8	293431	84	PFMAL13P4	AL049181 Plasmid
c 25	46	2.8	192676	61	AC010888	AC010888 Homo sapi
c 26	45.8	2.8	156550	60	AC015830	AC015830 Homo sapi
c 27	45.4	2.8	1063	53	CNS07A2Y	AL436084 T3 end of
c 28	45.2	2.8	2538	14	PSXNAGEN	X91858 Ptomycetes s
c 29	45.2	2.8	53932	68	AC023371	AC023371 Homo sapi
c 30	45.2	2.8	78379	12	AB006696	AB006696 Arabidops
c 31	45.2	2.8	160877	70	AC026560	AC026560 Homo sapi
c 32	45.2	2.8	174303	74	AC069506	AC069506 Homo sapi
c 33	45	2.8	5560	96	PFAL132422	AJ132422 Plasmid
c 34	44.8	2.8	5917	96	PFANT195	X02919 Plasmid
c 35	44.8	2.8	164399	96	PFMAL3P6	Z98551 Plasmid
c 36	44.6	2.7	35928	6	CELFALH10	U61954 Caenorhabdi
c 37	44.6	2.7	119918	82	AP000713	AP000713 Homo sapi
c 38	44.6	2.7	173805	91	CNS05TBW	AL352977 Human chr
c 39	44.6	2.7	293431	84	PFMAL13P4	AL049181 Plasmid
c 40	44.4	2.7	897	53	CNS07ABZ	AL436389 T7 end of
c 41	44.4	2.7	1032	53	CNS06LOP	AL404447 T7 end of
c 42	44.4	2.7	110000	84	PFMAL13P2_1	Continuation (2 of
c 43	44.4	2.7	110000	84	PFMAL13P2_2	Continuation (3 of
c 44	44.2	2.7	82360	13	AP000606	AP000606 Arabidops
c 45	44.2	2.7	341064	75	AC074226	AC074226 Arabidops

ALIGNMENTS

RESULT	1
E12754	
LOCUS	E12754 1622 bp DNA 24-JUN-1998
DEFINITION	Gentianatriflora mRNA for acyltransferase, complete cds.
ACCESSION	E12754
VERSION	E12754.1 GI:3251586
KEYWORDS	JP 1997070290-A/2.
SOURCE	unidentified.
ORGANISM	unidentified.
REFERENCE	1 (bases 1 to 1622)
AUTHORS	Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y., Yonekura, K., Mizutani, M. and Kusumi, T.
TITLE	GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL	Patent: JP 1997070290-A 2 18-MAR-1997; SUNTORY LTD
COMMENT	OS Gentianatriflora PN JP 1997070290-A/2 PD 18-MAR-1997 PR 30-JAN-1996 JP 1996046334 PF 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI PI ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHIRO, PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PC C12N15/09, A01H1/00, C07H21/04, C07K14/42, C12N9/10, (C12N9/10, PC C12R1:865), PC (C12N9/10, C12R1:19); CC strandedness: Double; CC topology: Linear; FH Key Location/Qualifiers FH

Db	781	AGCCACCTTCATCTACACCCCTATTGATATATCATGAAGCTCAAGAAATTCATTTCGTCAA	840
Qy	841	AAATCGCAACTTAACCGGTAGTAGTAATATATACTGTCAACTTTTCACGGTGACATCTGC	900
Db	841	AAATCGCAACTTAACCGGTAGTAGTAATATATACTGTCAACTTTTCACGGTGACATCTGC	900
Qy	901	ACTGATCTGGACATCCTTGTTCGAAATCATATTAGACACCGTCGTAAGAGAGAGGTTGGAAGA	960
Db	901	ACTGATCTGGACATCCTTGTTCGAAATCATATTAGACACCGTCGTAAGAGAGAGGTTGGAAGA	960
Qy	961	GGATAAATCATGAGCAAACTTATGTCTTTCATCAACTGCCGACAACGTTTTCGTCGCC	1020
Db	961	GGATAAATCATGAGCAAACTTATGTCTTTCATCAACTGCCGACAACGTTTTCGTCGCC	1020
Qy	1021	GATACCTCAAAATTAATTTGGAAATTCGATAGTGCCTTGTATGTTGGATCGCACTCATGA	1080
Db	1021	GATACCTCAAAATTAATTTGGAAATTCGATAGTGCCTTGTATGTTGGATCGCACTCATGA	1080
Qy	1081	GCACTTGTAGGAATAGAGGTTCTCGGTAGCTGCAACCGCCATCGGAGATGCTATCCA	1140
Db	1081	GCACTTGTAGGAATAGAGGTTCTCGGTAGCTGCAACCGCCATCGGAGATGCTATCCA	1140
Qy	1141	TAAGAGTTTACATGACTACGAAGGAATCTGAGAGAGATTCGGATATCGCCGCCGATC	1200
Db	1141	TAAGAGTTTACATGACTACGAAGGAATCTGAGAGAGATTCGGATATCGCCGCCGATC	1200
Qy	1201	AACATCTCGGCGACCAAGTTCGACGCTCATTTATCTGCTGGATCGGCACACGCAATGT	1260
Db	1201	AACATCTCGGCGACCAAGTTCGACGCTCATTTATCTGCTGGATCGGCACACGCAATGT	1260
Qy	1261	GCATGATTTTGATGCAGATTTTGGTTGGGAAAGCTTGAAAAGCATGAATCTGTTTCAAC	1320
Db	1261	GCATGATTTTGATGCAGATTTTGGTTGGGAAAGCTTGAAAAGCATGAATCTGTTTCAAC	1320
Qy	1321	TAATCCTCGGCAACACTAATTTTGTATCTCTCGGTCCAGAAATTTAAAGGAGCACTTGA	1380
Db	1321	TAATCCTCGGCAACACTAATTTTGTATCTCTCGGTCCAGAAATTTAAAGGAGCACTTGA	1380
Qy	1381	GCTTGGCATTTCTTTCCTTGAAGATAGGATGGACCATTTGCCACCATTTTACGAATTT	1440
Db	1381	GCTTGGCATTTCTTTCCTTGAAGATAGGATGGACCATTTGCCACCATTTTACGAATTT	1440
Qy	1441	CATCAATAGTCTCCATGTGAGGAGCCCTTGTGAAGAAAAAGTGTATCAATGTATAAAA	1500
Db	1441	CATCAATAGTCTCCATGTGAGGAGCCCTTGTGAAGAAAAAGTGTATCAATGTATAAAA	1500
Qy	1501	AGACAGACAAGTTATGATGCAACAATTTGTTTAGGAGATTACAATCCATCGGGAAGATG	1560
Db	1501	AGACAGACAAGTTATGATGCAACAATTTGTTTAGGAGATTACAATCCATCGGGAAGATG	1560
Qy	1561	TATCAACTCATCTCTCTATATATATATTCAAATTTGTTTAAAAAATAAAAAA	1620
Db	1561	TATCAACTCATCTCTCTATATATATATTCAAATTTGTTTAAAAAATAAAAAA	1620
Qy	1621	AA 1622	
Db	1621	AA 1622	
RESULT 2			
LOCUS	AB026494	1622 bp	mrna
DEFINITION	Gentiana triflora GAT106 mRNA for acyltransferase homolog, complete cds	PLN	04-APR-2000
ACCESSION	AB026494		
VERSION	AB026494.1	GI:7415596	
KEYWORDS	acyltransferase homolog.		
SOURCE	Gentiana triflora cDNA		
ORGANISM	Gentiana triflora		
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.		
REFERENCE	1 (sites)		
AUTHORS	Yonekura-Sakakibara,K., Tanaka,Y., Fukuchi-Mizutani,M., Fujiwara,H., Fukui,Y., Toshihiko,A., Yamaguchi,M. and Kusumi,T.		
TITLE	Molecular cloning and biochemical characterization of hydroxycinnamoyl-CoA:anthocyanin 3-O-glucoside-6-O-hydroxycinnamoyltransferase from Perilla frutescens and diverse plant acyltransferase homologs unpublished (1999)		
JOURNAL	2 (bases 1 to 1622)		
REFERENCE	Tanaka,Y. and Yonekura-Sakakibara,K.		
AUTHORS	Direct Submission		
TITLE	Submitted (21-APR-1999) to the DDBJ/EMBL/GenBank databases.		
JOURNAL	Yoshikazu Tanaka, Suntory Ltd., Institute for Fundamental Research; Wakayama-dai 1-1-1, Shimamoto, Osaka 618-8503, Japan (E-mail:Yoshikazu.Tanaka@suntory.co.jp, Tel:81-75-962-8807, Fax:81-75-962-8262)		
FEATURES	Location/Qualifiers		
source	1..1622		
	/organism="Gentiana triflora"		
gene	/db_xref="taxon:55190"		
	35..1474		
CDS	/gene="GAT106"		
	35..1474		
	/gene="GAT106"		
	/codon_start=1		
	/product="acyltransferase homolog"		
	/protein_id="BAA93452.1"		
	/db_xref="GI:7415597"		
	/translation="MAGSEDIKVLKCRVAPPDVAEFTVPLSFFDMRWLISDAEH		
	HLVRYRFPICPNKSFIISSIKSSLVLRHPLAGNLIMPVSSSDRMPEURKKGD		
	SVSTIAEESNDFYLAGDHQDSYKFNLDLPOLPEIVTSGDEVPLFALQVTVFSN		
	TGICIGRNLHVLGDSFLHFNKLVLDKSGDSLKFLPLSSLPYDMYVVDPPFH		
	IRRTYNERKLLKSGQPTVLNPAISKDEVRATFILHIDIMKLKFKFISSKRNLTGS		
	SNYLSPTVTTSALITWCLSKSLDTVVRKEVEEDKHAANLCAFINCRORAPPYONY		
	FGNCIVPCNVGSTEQLVGNBGLSVAATAIGDAIHKRHLDYEGILGRDWISPPRSTSA		
	APRSTLIYVGSQRNVHDFADFGKLEHESVSTNPSATLILISRRRFRKGALEL		
	GISLPKRNDAFAITFTNFINSLHVRSPL"		
BASE COUNT	478 a	345 c	322 g 477 t
ORIGIN			
Query Match	100.0%;	Score 1622;	DB 12; Length 1622;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1622;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	GAACCATCAATCCAAATTAATCTGATTTATTAAAGATGGCAGGAAATTCGAGGATATCAA	60
Db	1	GAACCATCAATCCAAATTAATCTGATTTATTAAAGATGGCAGGAAATTCGAGGATATCAA	60
Qy	61	AGTTCTTGAGAAATGCCGTGTTGCGCCACACCGGACGCGTCCGCGAGTTTACAGTCCC	120
Db	61	AGTTCTTGAGAAATGCCGTGTTGCGCCACACCGGACGCGTCCGCGAGTTTACAGTCCC	120
Qy	121	ACTGTGCTTTTCGACATGCGATGGTTGATCTCTGATGCAGAACCATCTCGATTTCTTA	180
Db	121	ACTGTGCTTTTCGACATGCGATGGTTGATCTCTGATGCAGAACCATCTCGATTTCTTA	180
Qy	181	CAGATTCGCCCATCTTCTCCAACTCTAAATTTATCATTTTCATTAATAATCGTCCCT	240
Db	181	CAGATTCGCCCATCTTCTCCAACTCTAAATTTATCATTTTCATTAATAATCGTCCCT	240
Qy	241	TTCCCTTGTCTCAAAACACTTTTCCGTTAGCCGGAAATTTGATTTGCCCGGTAGATTTC	300
Db	241	TTCCCTTGTCTCAAAACACTTTTCCGTTAGCCGGAAATTTGATTTGCCCGGTAGATTTC	300
Qy	301	CTCCGATAGATGCGGAGTTCGGTTTACAAGAAAGGGGACTCCGTTCTTTTAACAATTGC	360
Db	301	CTCCGATAGATGCGGAGTTCGGTTTACAAGAAAGGGGACTCCGTTCTTTTAACAATTGC	360
Qy	361	AGAATCGACATGGATTTTGTATCTCCCGGAGATCATCAGAGGATCTTATAAATTT	420
Db	361	AGAATCGACATGGATTTTGTATCTCCCGGAGATCATCAGAGGATCTTATAAATTT	420
Qy	421	CAACGATTTGATTCGCGAGCTGCCAGAACCGATTTGTAACCTCCGCGGAGAGTATTACC	480

|||||
 Db 421 CAACGATTTCATTCGCGAGTGCAGAACCGATTGTACCTCCGCGACGAAGTATTACC 480
 QY 481 ACTTTTGTCTTACAGGTGAGGTTCTCCCAACACCGGTATATGCAATGGACCAATCT 540
 Db 481 ACTTTTGTCTTACAGGTGAGGTTCTCCCAACACCGGTATATGCAATGGACCAATCT 540
 QY 541 TCATCAAGTCTTGGTGATGCGAGTCTTTTCTGCAATTTTAAATAATATGGGTTTGGT 600
 Db 541 TCATCAAGTCTTGGTGATGCGAGTCTTTTCTGCAATTTTAAATAATATGGGTTTGGT 600
 QY 601 TGCAAAATCCAAATGGAGATTCATTAAAGTTCCTTCCACTTTCTTCTTACCTATGACGA 660
 Db 601 TGCAAAATCCAAATGGAGATTCATTAAAGTTCCTTCCACTTTCTTCTTACCTATGACGA 660
 QY 661 CAGATCTGTGTGCAAGATCCATTTCTATTCGTCGAAATCTACAAATGAAGAAACT 720
 Db 661 CAGATCTGTGTGCAAGATCCATTTCTATTCGTCGAAATCTACAAATGAAGAAACT 720
 QY 721 GCTCAAAATCTCAGGCGACACCTACTGTCTTAAATCCAGCAATTTCTAAAGATGAAGTTCG 780
 Db 721 GCTCAAAATCTCAGGCGACACCTACTGTCTTAAATCCAGCAATTTCTAAAGATGAAGTTCG 780
 QY 781 AGCCACCTTCATCTACACCTATTTGATATCATGAAAGTCAAGAAATTCATTTGCTCAAA 840
 Db 781 AGCCACCTTCATCTACACCTATTTGATATCATGAAAGTCAAGAAATTCATTTGCTCAAA 840
 QY 841 AAATCGGAACCTTAACCGGTAGTACTTAATTAATCTGCACTTTCAGGTGACATCTGC 900
 Db 841 AAATCGGAACCTTAACCGGTAGTACTTAATTAATCTGCACTTTCAGGTGACATCTGC 900
 QY 901 ACTGATCTGGACATGCTTGTGAAATCATTTAGACACCGTCTGTAAGAGAGAGGTGGAAGA 960
 Db 901 ACTGATCTGGACATGCTTGTGAAATCATTTAGACACCGTCTGTAAGAGAGAGGTGGAAGA 960
 QY 961 GGATAAACATGCGACAACTTATGCTTTCATCACTGCGGACCAAGTTCGTCGCGC 1020
 Db 961 GGATAAACATGCGACAACTTATGCTTTCATCACTGCGGACCAAGTTCGTCGCGC 1020
 QY 1021 GATACCTCAAAATTTACTTTGGAAATTCATAGTGCCTTGTATGCGGATCGACTCATGA 1080
 Db 1021 GATACCTCAAAATTTACTTTGGAAATTCATAGTGCCTTGTATGCGGATCGACTCATGA 1080
 QY 1081 GCAACTTTGTAGAAATGAAGGTTTGTGCGTAGCTGCAACCGCCATCGAGATGCTATCCA 1140
 Db 1081 GCAACTTTGTAGAAATGAAGGTTTGTGCGTAGCTGCAACCGCCATCGAGATGCTATCCA 1140
 QY 1141 TAAGAGGTTACATGACTACGAAGAAATCTGAGAGGAGATTTGCGGCCCGCATC 1200
 Db 1141 TAAGAGGTTACATGACTACGAAGAAATCTGAGAGGAGATTTGCGGCCCGCATC 1200
 QY 1201 AACATCTGCGGCACCAAGGTCGAGCGCTCATTTATGCTGTTGGATCCGCACACCAATGT 1260
 Db 1201 AACATCTGCGGCACCAAGGTCGAGCGCTCATTTATGCTGTTGGATCCGCACACCAATGT 1260
 QY 1261 GCATGATTTTGCATGACGATTTGGTGGGAAAGCTTGAAGAGCATGAATCTGTTTCAAC 1320
 Db 1261 GCATGATTTTGCATGACGATTTGGTGGGAAAGCTTGAAGAGCATGAATCTGTTTCAAC 1320
 QY 1321 TAATCTTCCGCAACACTAATTTTGTATCTCTCGGTCCAGAGATTTAAAGGAGCAGCTTGA 1380
 Db 1321 TAATCTTCCGCAACACTAATTTTGTATCTCTCGGTCCAGAGATTTAAAGGAGCAGCTTGA 1380
 QY 1381 GCTTGGCATTTCTTGGCTAAGATAGATGAGGAGGATTTGCCACCAATTTTACGAATTT 1440
 Db 1381 GCTTGGCATTTCTTGGCTAAGATAGATGAGGAGGATTTGCCACCAATTTTACGAATTT 1440
 QY 1441 CATCAATAGTCTCCATGAGGAGCCCTTTGTAAGAAAGTGGTATCAATGTATATAAA 1500
 Db 1441 CATCAATAGTCTCCATGAGGAGCCCTTTGTAAGAAAGTGGTATCAATGTATATAAA 1500
 QY 1501 AAGACAGCAAGTTATGATGCAACAAATGTTTATAGGAGATTTACAAATCCATGGAAGATG 1560
 |||||

Db 1501 AAGACAGCAAGTTATGATGCAACAAATGTTTATAGAGATTACAAATCCATGGGAGATG 1560
 QY 1561 TATCAAACTCATCTCTAT 1620
 Db 1561 TATCAAACTCATCTCTAT 1620
 QY 1621 AA 1622
 Db 1621 AA 1622
 RESULT 3
 E12756
 LOCUS
 DEFINITION Perilla ocimoides mRNA for acyltransferase, partial cds.
 ACCESSION E12756
 VERSION E12756.1 GI:3251588
 KEYWORDS JP 1997070290-A/4.
 SOURCE unidentified.
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 1479)
 AUTHORS Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y.,
 Yonekura, K., Mizutani, M. and Kusumi, T.
 TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
 JOURNAL Patent: JP 1997070290-A 4 18-MAR-1997;
 SUNTORY LTD
 COMMENT OS Perilla ocimoides
 PN JP 1997070290-A/4
 PD 18-MAR-1997
 PF 30-JAN-1996 JP 1996046534
 PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
 ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
 MASAHITO,
 FUJIKU YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAHIKI PC
 C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
 C12R1:865),
 PC (C12N9/10,C12R1:19):
 CC strandedness: Double;
 CC topology: Linear;
 FH key Location/Qualifiers
 FT source 1..1479
 FT /organism='Perilla ocimoides' FT
 FT /tissue_type='leaves',
 FT /clone='pSAT208'
 FT CDS 1..1343
 FT /product='acyltransferase'.
 FEATURES
 source Location/Qualifiers
 1..1479 /organism='unidentified'
 /db_xref='taxon:32644'
 BASE COUNT 420 a 316 c 331 g 411 t 1 others
 ORIGIN
 Query Match 13.4%; Score 218; DB 10; Length 1479;
 Best Local Similarity 50.8%; Pred. No. 4.6e-44;
 Matches 741; Conservative 1; Mismatches 661; Indels 57; Gaps 7;
 QY 62 GTTCTTTCAGAAAATGCCGTGTTGGCCACACCGCGCGCTGCGGAGTTTACAGTCCCA 121
 Db 3 GTGATCGAAACGTGTAGAGTTGGCCGCCGCCGCGGACTCGGTGGCGGAGCAATCGGTCCG 62
 QY 122 CTGTCGTTTTTCGACATGCGATGCTCTGTGATCGAGAACACCATCTGCAATTTCTAC 181
 Db 63 CTCACATTTCTCGACATGACGTGGCTGCATTTTCATCCCATGCTTCAGCTCCTCTTCTAC 122
 QY 182 AGATTCGCGCATCTTCTCCCAAC---TCTAAATTTATCATTCATCCATTAATTCCTCC 238
 Db 123 GAATTCCTCTGTTCCCAAGCAACATTTTTCAGATCCATCGTTCCTCAAAACATCAACATCT 182
 QY 239 CTTTCCCTTCTCTCAACACACTTTTCTTCCGTTCAGCCGGGAATTTGATTTGCGCCGTAGAT 298

Db 63 CTCACATCTTCGACATGACGTGCTGCATTTTCATCCATGCTTCAGCTCCTCTCTAC 122
Qy 182 AGATTCCGCCATCTGTCGCCAAC---TCTAAATTTATCATTTTCATCCATTAATTCGTC 238
Db 123 GAATCCCTGTTCCAGCAACATTTTTCAGATCCATCGTTCGTCAGAACTCAAAATCT 182
Qy 239 CTTTCCCTTGTCTCAACACTTCTTCCGTTAGCCGGAATTTGATTTGGCCGCTAGAT 298
Db 183 CTCCTAAACCTCATACACTTCTTCCCTCTCTCATGCAATTTAACTACCCCTCA--- 239
Qy 299 TCTCCGATAGATCCCGAGTTCGGTTTCAAGAAAGGGGACTCCGTTTCTTTTAAACAT 358
Db 240 TCCCGGAGAAATGCCGAGTTCGGTATCTATCCGCGGACTCGGTTTCTTTCAACATC 299
Qy 359 GCAGATCGAGCATGGATTTGATTTATCTCCGCGGAGATCATCAGAGGATTTCTTATAA 418
Db 300 CGAATCTACGAGGACTTCGATGATCTCGCGAATCTCGCGAATCTCCAGATCTCCGCTTAGG 359
Qy 419 TTCAACGATTTGATTCGCGAGCTGCCAGAACCGATTTGTAACCTCCGCGAGCAAGTATTA 478
Db 360 CTCACAACTTTGTCCTTAATTCGCCGCCATTTGCGAAGAAATCCGATAGAAACTCTTC 419
Qy 479 CCACTTTTGTCTTACAGTGCAGGTGTTCTTCAACACCGGTATATGATTTGGAGCAAT 538
Db 420 CAAGTTTTCGCGGTGCAGGTGACTCTTTTCCAGCGCGAGCGTCTGATTTGGAAATGCA 479
Qy 539 CTTTCATCAAGTCTTGTGTGATGCAGGTCTTCTTCTGCAATTTATAAATATATGTTG 598
Db 480 ACGATCACACGGTTAGCGGCGCGCGTCTTCTCGGGTTTAAACGGCTTGGTCTTCA 539
Qy 599 GTTGACAAATCCAATGGAGATTCATTAAGGTCTCTTCCACTTTCTTCTCTACCTATGATC 658
Db 540 ATGACCAACACATTTGAATGAAGATCAAGATCAAGAAATTAATCTTCCAGTTTTC 599
Qy 659 GACAGATCTGTTGTCAGATTCATTTCAATTCATTCGTAAGAAATTCACAAATGAAGAAA 718
Db 600 GATAGATCCGTATATAATATCCGACGAAATTT-----GACTCCATTTATTTGGAGAAC 653
Qy 719 CTGCTCAATCTCAGGCGACACCTACTCTTCTTAATCCAGCAATTTCTAAGATGAAT 778
Db 654 CGCTAAATTTCTTTGCAATCTCGT-----CATCCTCATACCGACGCGCGAT 707
Qy 779 CGACCACTTTCATCCACACCTTATGATATCATGAGCTCAAGAAATTCATTTCTGTC 838
Db 708 CGAACAGTTTCGTTTACCCCAATCCAAATTAAGAAATTAAGGTTGATTCAGTCC 767
Qy 839 AAAATCGCAACTTAACCGGTAGTAGTAATTAATCTGTCAACTTTTCCAGGTGACATCT 898
Db 768 AGAGTTCCAAAGTTTAGTC-----CATCTCTCATCTTTTGTAGCGATTGCA 812
Qy 899 GCACTGATCTGGACATGCTTCTGAAATCATTTAGACACCGCTGTAAGAGAGAGAGTGAA 958
Db 813 GCTTATATGCTGGTGGCATTAACCAATATCTC-----ACAGCATGAGAGC 860
Qy 959 GAGGATAAATGACGAGCAAACTTATGTCTTTCATCACTGCCGACACGTTTGTCTCCG 1018
Db 861 CAAGACACGAGGATGCAATTTTCTTGATCCGCTGATCTAAGCCACGATTAGATCCG 920
Qy 1019 CCGATACCTCAAAATTCATTTGGAAATTCATAGTGCCTTGTATGTTGGGATTCAGCTCAT 1078
Db 921 CCGTTCTCTGAAATATCTTCGGAACTGCTTATCTGATCCGCGCTGCCGAGAAATCGCGG 980
Qy 1079 GAGCAACTGTGAGAAATGAAGGTGTGCGTAGCTGCTCAACCGCATCGGAGATGCTATC 1138
Db 981 CGAGAGCTGGGAGAGAAAGGGTGTCTTCCAGCTGAGTAATCGCGCGGAGATA 1040
Qy 1139 CATAAGAGTTTACATGACTACGAGGAATTTCTGAGGAGATTTGGATATCGCCCGCCGA 1198
Db 1041 AAAAAAGGATCAACGACAGAGAAATATTAGAAACGGTGGAGAAATGTCGCCCGAGATT 1100
Qy 1199 TCAACATCTCGCGGACCAAGTCAACGCTCATTTATGCTGTTGATCGGCACACGCAAT 1258

Db 1101 CGTAAAGC-----GTTGCAGAAATCATATTTTTCGGTGGCAGGATCGAGCAAG 1148
Qy 1259 GTGCATGATTTTGTGATGAGATTTTGGTGGCAAGCTTGAAGAGCATGATCTGTTTCA 1318
Db 1149 CTAGATCTTTACGTTGATGATTTTGGATGGGAGGCGGAGGAGCAAGAAATATTGTCG 1208
Qy 1319 ACTAATCCTTTCGGCAACACTAATTTTGTATCTCTCGTCCGTCAGAAAGATTTTAAAGGAGCACT 1378
Db 1209 ATTGATGGGAGAAATATGCAATGACACTTTTGAAGCCAGGATTTTGAAGAGGAGGATTG 1268
Qy 1379 GAGCTTGGCATTTCTTTCCTTAAAGATAGGATGACGATTTTGCACCATTTTACGAA 1438
Db 1269 GAGTTTGTCTTCTTTCCTTAAAGCAAAATGATGCTTTTGTCTTATTTTTCACGT 1328
Qy 1439 TTTCATCAATGCT 1450
Db 1329 GGAATTAATGCT 1340

RESULT 5
AB010708 1679 bp mRNA PLN 20-FEB-1999
LOCUS Gentiana triflora mRNA for Anthocyanin 5-aromatic acyltransferase,
DEFINITION complete cds.
ACCESSION AB010708
VERSION AB010708.1 GI:4185598
KEYWORDS Anthocyanin 5-aromatic acyltransferase.
SOURCE Gentiana triflora petal cDNA to mRNA, clone:pgAT4.
ORGANISM Gentiana triflora
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.
1 (bases 1 to 1679)
Fujiwara,H., Tanaka,Y. and Kusumi,T.
Direct Submission
Submitted (22-JAN-1998) to the DDBJ/EMBL/GenBank databases.
Hiroaki Fujiwara, Sutoryi LTD., Inst. of Fundamental Research;
1-1-1, Wakayamadai, Shamamoto-cho, Osaka 618-0024, Japan
(E-mail:Hiroyuki.Fujiwara@suntory.co.jp, Tel:+81-75-962-8807,
Fax:+81-75-962-8262)
2 (sites)
Fujiwara,H., Tanaka,Y., Yonekura-Sakakibara,K.,
Fukuchi-Mizutani,M., Nakao,M., Fukui,Y., Yamaguchi,M., Ashikari,T.
and Kusumi,T.
cDNA cloning, gene expression and subcellular localization of
anthocyanin 5-aromatic acyltransferase from Gentiana triflora
Plant J. 16 (4), 421-431 (1998)
99097837 Location/Qualifiers
1.1679
/organism="Gentiana triflora"
/db_xref="taxon:55190"
/clone="pgAT4"
/tissue_type="petal"
6.1415
/codon_start=1
/product="Anthocyanin 5-aromatic acyltransferase"
/protein_id="BAA74428.1"
/db_xref="GI:4185598"
/translation="MEQIQMKVLEKCVTPPSDTPDVELSLPVTFFDIPWLHLNKMQ
SLIFYDFPYPRTHFDIVIPNLKASLSLTLKHYPLSGNLLMPKSGEMPKFQYSRDE
GDSITITVAESDQDFYLGKQLVDSNDLHGLFVMPRVITRMDQYKVIPLVAVQVTV
FPMRGTAVALTAHSDAKSFVFNANAY INKFRDADLLSANLLPSFDRSTIKDL
YGLEETFWNEMODVLEMFSGSKPPPKVRATYVLSLAEIOKLNKVLNLRGSEPT
IRVTTTTCGYVWTCMVSKDDVSEESNDENELEYEFTADCGLLTLPCCPNYF
SKFLACVAKATKELVGDGLLVAAVAAIGEAIEKRLHNKNEKVLADAKTWSLNGIP
KIHMDAFAKIFEEGFCSLs"
1679
/notes="24 a nucleotides"
polyA_site 488 a 350 c 359 g 482 t
BASE COUNT
ORIGIN

Query Match		11.7%	Score 189.2;	DB 12;	Length 1679;
Best Local Similarity		52.2%;	Pred. No. 7.8e-37;		
Matches 728;		Conservative 0;	Mismatches 618;	Indels 48;	Gaps 12;
QY	57	TCAAAGTCTTCTGAGAAATCCGCTGTTGCCACACCGGAC---	CGCGTCGCCGAGTTTA	113	
DB	25	TGAAGGTCTTCTGAAANAATGCCAAGTTACACCAACCATCTGCACACACAGATGTCGAGTTAT	84		
QY	114	CAGTCCCACTGCTGTTTTCGAGATCGGATGGTTGATCTCTGATGCAGAACACCATCTCG	173		
DB	85	CGCTACCGGTAAACATCTTTCGATATCCCTGGTTGCACCTTGAATAAGATGCGAGTCCCTTC	144		
QY	174	ATTCTTACAGATTCGCCCACTCTTGTCGCAA---CTCTAAATTTATCATTTTCATCCATTA	230		
DB	145	TGTTTTACCACTTTCGTAACCAAGAACACATTTCTTGGACACTGTATCCCTAATCTTA	204		
QY	231	AATCGTCCCTTTCCTCTCTCAACACTTTCTTCGTTAGCCGGAAATTTGATTTGCG	290		
DB	205	AGGCTCTTTGCTCTCACTCTAAACACTAGCTCCGCTTAGCGGAAATTTGTTGATGC	264		
QY	291	CGGTAGATTCCTCCGATAGAAATGCCGAGTTGGTTAC-----AGAAGGGGACTCCG	344		
DB	265	CGATCAAAATCGGGCGA---AATGCCGAAGTTTCACTTCCGTAACCTCCGCTGATGAGGGGACTCGA	321		
QY	345	TTCTTTTAACAATTCGAGATCGAGCATGGATTTGATTTATCTCGCGCGAGATCATCAGA	404		
DB	322	TAACTTTGATCGTTGGGAGCTGACCAAGATTTTGACCTAACCTTAAGGTCACTCAACTGG	381		
QY	405	GGGATTTCTTATAAATTAACAGATTTGATTTCCGACGCTGCCAGAACCGATTTGTAACC---	461		
DB	382	TAGATTTCAATGATTTGCGTGGCTTTTATGTTATGTCACCGGGTTATAAGGACCATGC	441		
QY	462	CGGGGACCAAGTATTAACACATTTTTCCTTTTACAGGTGAGGTTCTTCCACACCGGTA	521		
DB	442	AAGACTATAAAGTGATCCCGCTCGTAGCGGTGCAAGTAACCGTTTTCCTCAACCGTGGCA	501		
QY	522	TATGCAATTCGAGCAATCTTCAATAGTTCTTGGTGATGCCAGTCTTTTCTGTCATTTTA	581		
DB	502	TAGCCGTGCTCTGACGGCACATCATTCATTTGCAGATGCTAAAGATTTTGTGAATGTTCA	561		
QY	582	ATAAATATGGGTTTGGTTGACAAATCAATGAGGATTCATTAAGTTTCTTCCACTTT	641		
DB	562	TCAATGCTTGGGCTATATAACAAT---TTGGGAAAGACGGGACTTGTGTCGCGCA	618		
QY	642	CTTCTCTACTATGTACGACAGATCTGGTGCAAGATCCATTTTCATATCTCGTCAAAA	701		
DB	619	ATCTTCTTCCATCTTTCGATAGATGATTAATCAAGATCTGTATGCCCTAGAGGAAACAT	678		
QY	702	TCTACAATGAAAGAAAATGCTCAAAATCTCAGGGCACACCTACTGTCTTAAATCCAGCAA	761		
DB	679	TTTGGAAACCAATGCAAGATGT---TCTTGAATGTTCTCTAGATTTGGAAGCAACCCC	735		
QY	762	TTTCTAAAGATGAAGTTTCGAGCACCTTCATCTACACCCCTATGTATATCATGAAGCTCA	821		
DB	736	CTCGATTCAACAAGGTACGAGCTACATATGTCTCTCCCTTGTGTAATCCAGAACTAA	795		
QY	822	AGAAATTCATTTTGGTCAAAAATCGCAACTTAACCGGTAGTAGTAATTAATCTGTCAA	881		
DB	796	AG-AACAAGTACTGAATCTCAGAGGATCCGAACCCACAATA-----CGGTAAACGA	846		
QY	882	CTTTCACGGTGACATCTGCACATCTGACATGCTTGTGCAANTCATTTAGAACCGCTCG	941		
DB	847	CGTTCACATGACGTGGATAGCTATGACATGCAATGATGTCATTAAGATGACGTTCG	906		
QY	942	TAAAGAGAAA---GGTGGAGAGGATAAACAATGCAGCAAACTTATGTCTTTCATCAACT	998		
DB	907	TATCAGAGGAATCATCAACGAGCAAAATGAGCTCGAGTCTTCACTTTTACAGCGGATT	966		
QY	999	GCGGACAAAGTTTGTCTCCGCCGATACCTCAAAATTAACCTTTTGAANAATGATGCTT	1058		
DB	967	GCGGAGGACTTCTGAGCGCCCCGTGTCGCGCTAACTACTTTTGGCAACTGTCTTGCGTCAT	1026		

QY	1059	GTATGTTGGGATCGACTCATGAGCAACTTTGTAGGAATGAAGGTTTTCGGTAGCTGCAA	1118		
DB	1027	GCCTTGCARAAACACACATATAAGAGTTAGTTGGGATATAAGGCTTCTTGTTCGATTCG	1086		
QY	1119	CGCCCATCGAGATGCTATCCATTAAGAGGTTTACATCACTACGAAAGAAATTTCTGAGAGG	1178		
DB	1087	CAGCTATTGGAGAAGCATTTGAAGAGGTTGCACACGAAAGGCTTCTTTCGAGATG	1146		
QY	1179	ATTGGATATCGCGCCCGCCGNTCAACATCTGCGGCACCAAGGTCGAGCTCATTTATGTCG	1238		
DB	1147	CAAAAACCTTTGGTTATCGGAATCTAAT-----GGAATCCCTTCAAAAAGATTTCTCG	1197		
QY	1239	TTGGATCCCAACCAACGCAATTTGTCATGATTTTTCATGATGATTTTGGTGGGAAAGCTTG	1298		
DB	1198	GGATTACCGGATCGCTTAAGTTCGATTCGTATGTTGATTTTGGATGGGAAAGCTTG	1257		
QY	1299	AAAGCATGAATCTGTTTCACTAATCTTTCGCGCAACACTAATTTTGTGATCTCTCGGTCCA	1358		
DB	1258	CAAAATTTGACA---TTAGTCTCTGTTGATATGCAGAAATTTGATTTATGTTGATTCAGTCCA	1314		
QY	1359	GAAGATTTAAAGGACACTTGAGCTTGGCATTTTTCCTTGCCTAAGAAATAGGATGACGCAT	1418		
DB	1315	GGATTTTGAAGAAAGTGTGGAGATTTGGAGTATCATTTGCCCTAAGATTCATATGGATGCAT	1374		
QY	1419	TTGCCACCACTTTT 1432			
DB	1375	TTGCAAAATCTTT 1388			
RESULT	6				
LOCUS	E12753	1703 bp	DNA	PAT	24-JUN-1998
DEFINITION	Gentianatriflora mRNA acyltransferase, complete cds.				
ACCESSION	E12753				
VERSION	E12753.1	GI:3251585			
KEYWORDS	JP 1997070290-A/1.				
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1703)				
AUTHORS	Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y., Yonekura, K., Mizutani, M. and Kusumi, T.				
TITLE	GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY				
JOURNAL	Patent: JP 1997070290-A 1 18-MAR-1997;				
COMMENT	SUNTORY LTD				
	OS	Gentianatriflora			
	PN	JP 1997070290-A/1			
	PD	18-MAR-1997			
	PF	30-JAN-1996	JP 1996046534		
	PR	17-FEB-1995	JP 95P 67159, 29-JUN-1995	JP 95P 196915	PI
	ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHIRO,				
	PI	FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI	PC		
	C12N15/09, A01H1/00, C07H21/04, C07K14/42, C12N9/10, (C12N9/10, PC C12R1:865),				
	PC	(C12N9/10, C12R1:19);			
	CC	Strandedness: Double;			
	CC	Topology: Linear;			
	FH	Key	Location/Qualifiers		
	FT	source	1. .1703		
	FT		/organism='Gentianatriflora'		
	FT		/variety='japonica'		
	FT		/tissue_type='petal'		
	FT		/clone='pgAT4'		
	FT	CDS	6. .1415		
	FT		/product='acyltransferase precursor'	FT	
	mat_peptide	9. .1412	/product='acyltransferase'		
FEATURES	Location/Qualifiers				
	source	1. .1703			
		/organism='unidentified'			
		/db_xref='taxon:32644'			

BASE COUNT 512 a 350 c 359 g 482 t

Query Match 11.7%; Score 189.2; DB 10; Length 1703;
Best Local Similarity 52.2%; Pred. No. 7.9e-37;
Matches 728; Conservative 0; Mismatches 618; Indels 48; Gaps 12;

QY 57 TCAAGTCTTTCAGAAATGCGGTGTGGCCACCACCCGGAC---GCCGTGCGCGAGTTTA 113
DB 25 TGAAGTCTTTCAGAAATGCGGTGTGGCCACCACCCGGAC---GCCGTGCGCGAGTTTA 84
QY 114 CAGTCCACATGTCGTTTTTTCAGATCGATGCGGTGTGATCTCTGATGCGAAGACCATCTGC 173
DB 85 CGTACCGGTAAATCTTCGATATCCCTGGTTCGACTTGAATGAAGATGCGAGTCCCTTC 144
QY 174 ATTTCTACAGATTCGGGCATCTTGTCCCAA---CTCTAAATTTATCATTTTCATCCATTA 230
DB 145 TGTTTTACGACTTTCCTGATCCCAAGAACACATTTCTTGGACACTGTTATCCCTAATCTTA 204
QY 231 AATCGTCCCTTTCCTTGTCTCAACACATTTCTTCCGTTAGCCGGGAATTTGATTTGGC 290
DB 205 AGGCTCTTGTCTCTCACTCTAAACACATCTAGTCCGCTAGCGGAATTTGTTGATGC 264
QY 291 CGGTAGATTCCTCCGATAGATGCGGAGTTGCGTTAC-----AAGAAAGGGGACTCCG 344
DB 265 CGATCAATCGGGCGA---AATCGGAAGTTTTCAGTACTCCCGTGAAGGGCGACTCGA 321
QY 345 TTTCTTAAACAAATTCAGATCGAGATGATTTTGTATTTATCTCCCGGAGATCATCAGA 404
DB 322 TAACCTTTGATGCTTTCGCGAGTCTGACGAGATTTTGTACTTAAAGGTCATCAACTGG 381
QY 405 GGGATTTCTTAAATTCACGATTTGATTCGCGAGTCCGACAGACCCGATTTGTAACC---T 461
DB 382 TAGATTCGAATGTTGATGCGCTTTTATGTTATGTCACCGGTTATAGGACCATGC 441
QY 462 CCGCGCAGAGATTAATACCACTTTTTCCTTACAGGTGAGCGTGTCTTCCCAACACCGGTA 521
DB 442 AAGACTATAAGTATCCGCTCGTAGCGGTGCAAGTAACCGTTTTCCTTAAACCGTGCA 501
QY 522 TATGATTTGAGCAATCTTCATCAAGTCTTGTGTGATGCGAGTCTTCTTCTGATTTTA 581
DB 502 TAGCGGTGCTTACCGGCATCATTTCAATTTGAGATGCTTAAAGTTTGTATGTTCA 561
QY 582 ATAAATATGTTGTTGTCACAAATCCAAATGAGATTTCAATTAAGTTCCTTCCACTTT 641
DB 562 TCAATGCTTGGCCTATATTAACAAAT---TTGGGAAGACCGGACTTGTGTCCCGCA 618
QY 642 CTTCTACCTATGACAGATGCTGTGTGCAAGATTCATTTATTCGTGCAAAAA 701
DB 619 ATCTTCTTCCATCTTTCGATAGATCGATAATCAAGATCTGTATGGCTAGAGGAACAT 678
QY 702 TCTACATGAAGAAACTGCTCAATCTCAGGCGACACCTACTGTTCTTAATCCAGCA 761
DB 679 TTTGGAACGAATGCAAGATGT---TCTTGAATTTGTTCTAGATTTTGGAGCAAAACCC 735
QY 762 TTTCTAAGATGAAGTTCGAGCCACCTTCACTCCACCCCTATTGATATCATGAAGTCA 821
DB 736 CTCGATTAACAAGTAGGAGTACATATGTCCTCTCCCTTGTGTAATCCAGAGTAA 795
QY 822 AGAATTTCTTTCGTAATAAATCGCAACTTAACCGGTAGTAGTAATTAATCTGTCAA 881
DB 796 AG-AACAAAGTACTGTAATCTCAGAGGATCCGACCGACAATA-----CGTGAACGA 846
QY 882 CTTTACGGGTGACATCTGCACTGATCTGGACATGCTTGTGGAATCATAGACCCGTCG 941
DB 847 CGTTCAAAATGACGTGTGGATACGATGAGATGAGATGCTGGTCAAAATCAAAAGATGACGTCG 906
QY 942 TAAGAGAGAA---GGTGAAGATGATAAATCATGACGCAAACTTATGTGCTTTTCAACT 998
DB 907 TATCAGAGGAATCATCAACGACGAAATAGAGCTCGAGTACTTTCAGTTTACACGGATT 966
QY 999 GCCGACAAAGTTTGTCTCCGCGGATACCTCAAAATTAATTTGGAATTTGCATAGTGCCTT 1058

Db 967 GCCGAGACTTCTGACGCCCCCGTGTCCGCTAACTACTTTGGCAACTGTCTTCGTCAT 1026
QY 1059 GTATGTTGGATCGACTGACGAGCAACTTGTAGCAATGAAGGTTGTGCGTACCTCAA 1118
Db 1027 GCGTTGCAAAAGCAACACATTAAGAGTTAGTTGGGATAAAGGGCTTCTTTGTCAGTTG 1086
QY 1119 CCGCATCGGAGATGCTATCATTAAGAGTTTACATGACTACGAAGGAATTTCTGAGAGAG 1178
Db 1087 CAGCTATTGGAGAGCCATTCAAAAGAGTTGCAACGAAGAGGCGCTTCTTCAGATG 1146
QY 1179 ATTGGATATCGCCGCCCGATCAACATCTGCGGCACCAAGTGCAGCGCTCATTTATGTCG 1238
Db 1147 CAAAAAATCTGTTTATCGGAATCTAAT-----GGAATCCCTTCAAAAAGATTTCTCG 1197
QY 1239 TTGATCGGCACACGCAATCTGCATGATTTTGATGCGAGATTTGGTTGGGAAAGCTTG 1298
Db 1198 GGATTACCGGATCGCTTAAAGTTGATGCTGATGTTGATGTTGGATGGGCAAGCGCTG 1257
QY 1299 AAAAGCATGAATCTGTTTCAACTAATCTTCGCAACACTAATTTTGATCTCTCGGTCCA 1358
Db 1258 CAAATTTGACA---TTACCTCTGTTGATGATGAGATTTGATTTGATGATTCAGTCCA 1314
QY 1359 GAAGATTTAAAGAGCAGCTTGTGATTTGCGATTTCTTTTGCCTTAAGATAGGATGAGCGCAT 1418
Db 1315 GGGATTTTGAAGAGGTGTGGAGATTTGGAGTATGATGTTGATGTTGATGATGATGATGAT 1374
QY 1419 TTGCCACCAATTTT 1432
Db 1375 TTGCAAAATCTTT 1388

RESULT 7
E12757
LOCUS E12757 1508 bp DNA PAT 24-JUN-1998
DEFINITION Senecio cruentus mRNA for acyltransferase,,partial cds.
ACCESSION E12757
VERSION E12757.1 GI:3251589
KEYWORDS JP 1997070290-A/5.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1508)
AUTHORS Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y.,
Yonekura,K., Mizutani,M. and Kusumi,T.
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 5 18-MAR-1997;
SUNTORY LTD
COMMENT OS Senecio cruentus
PN JP 1997070290-A/5
PD 18-MAR-1997
PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAARI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
C12R1:865),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1508
FT /organism='Senecio cruentus'
FT /tissue_type='petal.'
FT /clone='pCAT48'
FT /product='acyltransferase'.
FT CDS 1..1367
FT Location/Qualifiers
1..1508
/organism='unidentified'
/db_xref='taxon:32644'

BASE COUNT 442 a 293 c 296 g 477 t
ORIGIN

Query Match 8.6%; Score 139.8; DB 10; Length 1508;
Best Local Similarity 48.7%; Pred. No. 1.9e-24;
Matches 677; Conservative 0; Mismatches 642; Indels 72; Gaps 8;

QY 57 TCAAGCTCTTCGAGAAATCCGCTGTTGCGCCACACCGCGCGCGGAGTTTACAG 116
DB 1 TGAACATCTCGAACATGCGCGAATATCGCGCCCTCGCGCCACATCGCGT 60
QY 117 TCCACTGCTGTTTTCGACATCGCATGTTGATCTCTGATGCGAGAACACCATCTGCATT 176
DB 61 TATCTCTTACTTCTTCGACATTAATCTGCTACTCTCTCCCTCCGGTCCACCATCTTTTCT 120
QY 177 TCTACAGATTCCGCCATCTCTGTCCTCAACATCTTAATTTATCATTTTCATCA 233
DB 121 TCTATGACTTTCACATCTTCTAATCCATTTTCATGACACTAATTTGTTCCGACGCTAAAC 180
QY 234 CGTCCCTTTCCCTTCTCAACACTTTCTTCCGTTAGCGGGAATTTGATTTGGCCGG 293
DB 181 AATCTTTATCGGTCACTCTTCAACATTTTTCGGTTTGTAGTAATTTGATTTGATTTTC 240
QY 294 TAGATTCTCCG-----ATAGATGCCGGAGTTGCGTTTACAAGAAAGGGGACT 341
DB 241 CTAACTGATGTTGCGGTTTAAATAAACCAGAAATAAACACGTTTGAAGGTGATT 300
QY 342 CCGTTTCTTAACAATTGAGATCGAGATCGATGATTTTGTATATCTCGCGGAGATCATC 401
DB 301 CTGTTCTGTTACTTTTGCAGAAATGTTCTTGACTTTTAAATAATTTGACAGGAATCATC 360
QY 402 AGAGGATTCTTATAAATCAACGATTTGATTCGCGAGCTGCCAGAACCGATTGTAACCT 461
DB 361 CTCGAAATGTGAAACTTTTATCCACTTTGATCTTATGGAATGCAATCAATTTAT 420
QY 462 CCGCGCACCAAGTATTACCACCTTTTTCCTTACAGGTGACGGTGTCTCCACACCGGTA 521
DB 421 GTGATTGCGTCAGGTCCTCTTTTCACTTCAAGTGACGTTTTTTCGGGCTCGGTA 480
QY 522 TATGCAATTGACGCAATCTTCAAGTTCTTGGTGATGCGAGTTCTTTTCTGCATTTTA 581
DB 481 TATCACTAGGAATGACGAATCATCATGAGCTTGGTGACGCTAGCAGCGGTTCAACTTTT 540
QY 582 ATAAATATGTTGTTGTTGCAAAATCCAAATGAGATTCATTAAGTTCCCTCCACTTT 641
DB 541 TGAAGGGTGAGCTTCGATTTTCAATCTGGTGTAGT-----CGGT 582
QY 642 CTCTCTACTATGTACGACAGATCTGTGTCGAAGATCCATTTTCATATTCGTCGAAAA 701
DB 583 CTTTTTAAACGAAAGGATCTCCACCGTTTTTTCATAGATTGATTAACATCCACATTTAG 642
QY 702 TCTAATGAAGAAATGCTCAATCTCAGGGCACACTACTGTTCTAAATCCAGCAA 761
DB 643 ATGAAATTAAGTTGAGACATACAAGCTCGAAAGTTTATATAACCTTCGAGCCTGTGTG 702
QY 762 TTCTAAAGATGAAGTTCGAGCCACCTTCATCTACACCTATTGATATCATGAAGCTCA 821
DB 703 GTCCCACTGATAAAGTTTCGGTCAACCTTTGTTGACCCGAACTAATCAATCTACTAA 762
QY 822 AGAAATTCATTTGCTCAAAAAATTCGCAACTTTAACCGGTAGTAGTAATTAATCTGTCAA 881
DB 763 AGAAAGGCTTTAACCAAGTGCCAACTT-----GGAGTACATGTCAT 807
QY 882 CTTTCAGGTGACATCTGCACATGATCTGCAGATCTGTTGCGAAATCATTTAGACACCGTCG 941
DB 808 CTTTTACGGTAACCTTGTGTTATATATGAGTTGTCATAGCGAAATCACTCGTAAAAATAG 867
QY 942 TAGACAGAGGTTGGAAGAGGATAACATGTCAGCAAACTTATGCTGTTTCATCACTGCC 1001
DB 868 GAGAAAGAAAGGGCGAAGACG-----AGTTAGAACAGTTTCAATCACTCATTTGTTGC 921
QY 1002 GACAAGTTTGTCTCCGCGGATACCTCAAAATTTACTTTTGAAATTCATAGTGCCTTGTA 1061

DB 922 GATCTCGTCTTGATCCACCAATTCACAGCCCTACTTTGGTAACTGTGGTGCCACCATGTG 981
QY 1062 TGGTGGATCGCATCATGAGCAACTTTGTAGGAAATGAAGGTTTGTGGTGTGCGACCG 1121
DB 982 TCCGACCTTTAAAAAATGTCGTTTGTGACTAGCGAAAATGGGTATGCACTTGGTGTAAAG 1041
QY 1122 CCATCGGAGATGCTATCCATAAGAGGTTTACATGACTACGAGGAATTTCTCAGAGGAGATT 1181
DB 1042 TAATTGGAGAGTCTATATGCAAAATGATATATTAATAGGACGAATCTTGAAGATGCGC 1101
QY 1182 GGATATCGCGCCCGCCGATCAACATCTCGGACCAAGGTCGACGCTCATTTATGCTGTG 1241
DB 1102 CGAGATGGCATGAACCTTTCATGATCCCG-----CTAGGAAGATTGGTGTGCTG 1152
QY 1242 GATCCCAACACCAATGTCATGATTTTGCATGCAGATTTTGGTGGGAAAGCTTGAAA 1301
DB 1153 GTACACTAAGCTCAACTTGTACGACTTT-----GATTTGGTGGGAAAGCCGATAA 1206
QY 1302 AGCATGAATCTGTTTCAACTTAATCTTCCGCAACACTTAATTTTGTCTCTCGGTCCAGAA 1361
DB 1207 AGTATGAGACTGTTTCAATAGA---CTATAATAGCTCGATTTCTATAAATGCAAGCAAAA 1263
QY 1362 GATTTAAAGAGACATTTGAGCTTTGGCATTTCTTGGCTAAGAATAGGATGAGCGCATTTG 1421
DB 1264 CATCAGCAACAAGATCTTGRAATTTGGATTTGAGTCTACCGAGTATGCAAAATGGAGCGGTTT 1323
QY 1422 CCACCATTTT 1432
DB 1324 CTAGCATCTTT 1334

RESULT 8

AC005308/c

LOCUS

DEFINITION

AC005308

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

REFERENCE

AUTHORS

JOURNAL

COMMENT

On Nov 12, 2000 this sequence version replaced gi:8810450.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 178273: contig of 178273 bp in length.

Location/Qualifiers

1. 178273

/organism="Plasmodium falciparum"

/db_xref="taxon:5833"

/chromosome="12"

/clone="RFYAC492"

/clone="3D7"

BASE COUNT 71271 a 17228 c 17167 g 73607 t


```

QY 1596 TGTTTTAAAAA..... 1621
LOCUS   ||| ||||| ||||| |||
Db 72989 TATTATATAA.....TATA 72964

RESULT 11
AC004709/c
DEFINITION
Plasmodium falciparum chromosome 12, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.
AC004709
AC004709.3 GI:4558585
KEYWORDS HTG: HTGS-PHASE1.
SOURCE   malaria parasite P. falciparum.
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE
AUTHORS Hyman,R.W., Fung,E.L., Qin,F., Tamaki,T., Kurdi,O.B., Conway,A.B.
and Davis,R.W.
1 (bases 1 to 196149)
Plasmodium falciparum 3D7 chromosome 12
TITLE    Unpublished
JOURNAL
REFERENCE
AUTHORS Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.
DIRECT SUBMISSION
TITLE    Direct Submission
JOURNAL Submitted (21-MAY-1998) Stanford DNA Sequencing and Technology
Center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA
COMMENT  On Apr 2, 1999 this sequence version replaced gi:4337173.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 47631: contig of 47631 bp in length
* 47632 47831: gap of unknown length
* 47832 179129: contig of 131298 bp in length
* 179130 179329: gap of unknown length
* 179330 196149: contig of 16820 bp in length.
* Location/Qualifiers
* 1..196149
* /organism="Plasmodium falciparum"
* /db_xref="taxon:5833"
* /chromosome="12"

BASE COUNT 80057 a 19753 c 18800 g 77138 t 401 others
ORIGIN

Query Match 3.1%; Score 50; DB 60; Length 196149;
Best Local Similarity 58.9%; Pred. No. 0.12;
Matches 86; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 1476 AAAAAAGTGTATCAATGTATATAAAAAAGACAGACAAGTATGATCAACAATGTTTATAG 1535
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164454 AAATTAACCAACCAAAAAACCAAAAAATATAAATTCAAATTTATTAAACTTCAAT 164395

QY 1536 GAGATTACAATCCATGGGAAGATGATCAAACTCATCTCTCTATATATATATATTCAT 1595
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 164394 AAATGACAATAAATGTTAATAAGTTTTCACAAATTTGATATAAAAAATATAGTTAAT 164335

QY 1596 TGTTTTAAAAA..... 1621
Db 164334 TATTATATAA.....TATA 164309

RESULT 12
MISOLSUR
LOCUS   MISOLSUR 42781 bp DNA circular PLN 08-AUG-2000
DEFINITION Scenedesmus obliquus complete mitochondrial genome.
ACCESSION X17375 AJ271733 AJ272528 AJ277429 AJ400708

X17375.2 GI:7711033
16S ribosomal RNA; 16S rRNA gene; 23S ribosomal RNA; 23S rRNA gene;
atp6 gene; atp9 gene; ATPase subunit 6; ATPase subunit 9; cob gene;
cox1 gene; cox2 gene; cox3 gene; cytochrome b; cytochrome oxidase
subunit 1; cytochrome oxidase subunit 2; cytochrome oxidase subunit
3; nad1 gene; nad2 gene; nad3 gene; nad4 gene; nad4L gene; nad5
gene; nad6 gene; NADH dehydrogenase subunit 1; NADH dehydrogenase
subunit 2; NADH dehydrogenase subunit 3; NADH dehydrogenase subunit
4; NADH dehydrogenase subunit 4L; NADH dehydrogenase subunit 5;
NADH dehydrogenase subunit 6; orf130; orf148; ORF151; ORF367;
ORF390; orf76; ORF90; rnl1 gene; rnl2 gene; rnl3 gene; rnl4 gene;
rns1 gene; rns2 gene.
Scenedesmus obliquus.
SOURCE   Mitochondrion Scenedesmus obliquus
ORGANISM Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
Chlorococcales; Scenedesmaceae; Scenedesmus.
REFERENCE
AUTHORS Kuck,U., Godehardt,I. and Schmidt,U.
1 (bases 1 to 42781)
TITLE    A self-splicing group II intron in the mitochondrial large subunit
rRNA (LSURNA) gene of the eukaryotic alga Scenedesmus obliquus
JOURNAL Nucleic Acids Res. 18 (9), 2691-2697 (1990)
MEDLINE 90251444
REFERENCE
AUTHORS Winkler,M. and Kuck,U.
2 (bases 1 to 42781)
TITLE    The group IIB intron from the green alga Scenedesmus obliquus
mitochondrion: molecular characterization of the in vitro splicing
products
JOURNAL Curr. Genet. 26 (6), 495-502 (1991)
MEDLINE 92145774
REFERENCE
AUTHORS Kuck,U., Jekosch,K. and Holzamer,P.
3 (bases 1 to 42781)
TITLE    DNA sequence analysis of the complete mitochondrial genome of the
green alga scenedesmus obliquus: evidence for UAG being a leucine
and UCA being a non-sense codon
JOURNAL Gene 253 (1), 13-18 (2000)
MEDLINE 20384610
REFERENCE
AUTHORS Kueck,U.
4 (bases 1 to 42781)
TITLE    Direct Submission
JOURNAL Submitted (20-DEC-1989) Kueck U., Lehrstuhl fuer Allgemeine
Botanik, Ruhr-Universitaet Bochum, Postfach 10 21 48,
Universitaetsstr.150, D-44780 Bochum
REMARK   Revised by [4]
REFERENCE
AUTHORS Jekosch,K.
DIRECT SUBMISSION
TITLE    Submitted (05-MAY-2000) Jekosch K., Ruhr-Universitaet Bochum, LS
fuer Allgemeine und Molekulare Botanik, Universitaetsstr. 150,
D-44780 Bochum, GERMANY
COMMENT  On May 23, 2000 this sequence version replaced gi:7573226
gi:7160201 gi:7619793 gi:7573572 gi:13633.
FEATURES
Location/Qualifiers
1..42781
/organism="Scenedesmus obliquus"
/organelle="mitochondrion"
/strain="KS3-2"
/db_xref="taxon:3088"
/country="Germany"
complement(770..1000)
/gene="orf76"
complement(770..1000)
/gene="orf76"
/function="unknown"
/codon_start=1
/transl_table=22
/product="ORF76"
/db_xref="GI:7711034"
/translation="MCISFIQIKNFYFLSFIFLFIKNFYLSFIKNFYLSFIKNFYLSFI
KNLYNLLFKKCKDYKSDKSFCCIFYNIKKSL"
/gene="tRNA-Ile"
/product="transfer RNA Ile"

```



```

ACCESSION AF204057
VERSION AF204057.1 GI:8099190
KEYWORDS
SOURCE Scenedesmus obliquus.
ORGANISM Mitochondrion Scenedesmus obliquus
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
Chlorococcales; Scenedesmaceae; Scenedesmus.
REFERENCE 1 (bases 1 to 42919)
AUTHORS Nedelcu, A.M., Lee, R.W., Lemieux, C., Gray, M.W. and Burger, G.
TITLE The complete mitochondrial DNA sequence of scenedesmus obliquus
reflects an intermediate stage in the evolution of the green algal
mitochondrial genome
JOURNAL Genome Res. 10 (6), 819-831 (2000)
MEDLINE 20314873
REFERENCE 2 (bases 1 to 42919)
AUTHORS Burger, G.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1999) Departement de Biochimie, Universite de
Montreal, 2900 Boulevard Edouard-Montpetit, Montreal, Quebec H3T
1J4, Canada
FEATURES
source
1..42919
/organism="Scenedesmus obliquus"
/organelle="mitochondrion"
/strain="UTEX 78"
/db_xref="taxon:3088"
1..716
/gene="rnl_a"
/note="similar to Escherichia coli sequence"
/product="large subunit ribosomal RNA"
1..716
/gene="rnl_a"
786..858
/gene="trnA(ugc)"
/product="trnA-Ala"
/anticodon="(pos:819..821,aa:Ala)
786..858
/gene="trnA(ugc)"
join(859..1871,2480..2528,3802..4158)
/gene="rnl_d"
/note="similar to Escherichia coli sequence"
/product="large subunit ribosomal RNA"
859..4158
/gene="rnl_d"
859..1871
/gene="rnl_d"
/number=1
/number=1
/gene="rnl_d"
1872..2479
/gene="rnl_d"
/note="Group IIB; not present in Escherichia coli
sequence"
/number=1
2480..2528
/gene="rnl_d"
/number=2
2529..3801
/gene="rnl_d"
/note="Group IAL; not present in Escherichia coli
sequence"
/number=2
3027..3674
/gene="orf215"
/number=2
3027..3674
/gene="orf215"
/note="Orf inside an intron"
/codon_start=1
/transl_table=22
/product="orf215"
/protein_id="AAF72046.1"
/db_xref="GI:8099191"
/translacion="MLPPLFVEAIGNRIGDRSLYTTTKGSKIKLEQGRVKNKDYLYH
LPFLYKGMTNPKPYRIPIKVTGTYTRGVKISYGFRTIHPAFDKIYNFFTCNGKKT
YREGLITNHLTSVGLSYWVRDQSLQKNNEILRTMGFTQKENIQIVBELNRKFKLNS
KSKDFAC"
exon 3802..4158
/gene="rnl_d"
/number=3
4335..4407
/gene="trnH(gug)"
/product="trnA-His"
/anticodon="(pos:4368..4370,aa:His)
4335..4407
/gene="trnH(gug)"
4644..5018
/gene="rnl_c"
/note="similar to Escherichia coli sequence"
/product="large subunit ribosomal RNA"
4644..5018
/gene="rnl_c"
5084..5156
/gene="trnN(guu)"
/product="trnA-Asn"
/anticodon="(pos:5117..5119,aa:Asn)
5084..5156
/gene="trnN(guu)"
5396..5469
/gene="trnR(acy)"
/product="trnA-Arg"
/anticodon="(pos:5430..5432,aa:Arg)
5396..5469
/gene="trnR(acy)"
5525..5596
/gene="trnM(cau)"
/product="trnA-Met"
/anticodon="(pos:5557..5559,aa:Met)
5525..5596
/gene="trnM(cau)"
6066..6138
/gene="trnK(uuu)"
/product="trnA-Lys"
/anticodon="(pos:6099..6101,aa:Lys)
6066..6138
/gene="trnK(uuu)"
6448..7107
/gene="nad6"
6448..7107
/gene="nad6"
/EC_number="1.6.5.3"
/codon_start=1
/transl_table=22
/product="NADH dehydrogenase subunit 6"
/protein_id="AAF72047.1"
/db_xref="GI:8099192"
/translacion="MSSSFALITATGALMSVFLMMSKNPVHVSFYLVLFLHCSGLLY
LLGLEFYVLQLLYVYGALAILFLVWMLDIPSTEILRHQGTYPVAGVGLGLLILS
ALLAFQPLVESFPQTSPIPTSTFTGQNSDFSSRSRHSFNWNTRTATSTVANGIR
LYGVHADLIIRSLLLVAMVGAVGLTLKRPVSPNQDVFROHYVDFOKSLVQIPQNK
I"
7148..8347
/gene="cob"
7148..8347
/gene="cob"
/EC_number="1.10.2.2"
/codon_start=1
/transl_table=22
/product="apocytochrome b"
/protein_id="AAF72048.1"
/db_xref="GI:8099193"
/translacion="MKTSLKLRGSLWDOPMLAVQNHLSIYPTPSNLNGNYNMGVLA
GLCLVLIQRTGIFLAMYTAHVLYVWILLIMILTAFTGVLPWGQSLAGRTVITSLAS
MHLFRGLYTSVAOPREFVLYVWILLIMILTAFTGVLPWGQSLAGRTVITSLAS
ALPLYSTVGLWLGGEFVADNPTLNFYSFHYLFPFILLRSLHLRALHOVGSNPL
GNAQSDLVDFYPYPTTKDLVAALRLSFGAAYLVNFPBELLGHDPDNPANPYATPRH
IYWPYFWLVAYAILRSIENKLAGVRAIGLVFAALAAPTMHKKPSCRGFSFRRLHKVLV
WALIGDFMLLSYLGQOPVEPPYVVLGQRTVREFFSLILGLCVRWADALLANFSLKKS

```

[illegible][illegible]


```

LOCUS       AC023268      144457 bp      DNA      HTG      01-MAR-2000
DEFINITION  Homo sapiens chromosome 9 clone RP11-168D21 map 9, WORKING DRAFT
ACCESSION   AC023268
VERSION     AC023268.2  GI:7139920
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 144457)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
JOURNAL     1 (bases 1 to 144457)
REFERENCE   2 (bases 1 to 144457)
AUTHORS     Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
             Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
             Boguslavsky,L., Boukhgaiter,B., Brown,A., Burkett,G., Castile,A.,
             Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cooke,P.,
             Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
             Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
             Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
             Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
             Landers,T., Lehoczy,J., Levine,R., Liew,C., Liu,G., Locke,K.,
             Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
             McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
             Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
             Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
             Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
             Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
             Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye.W.J.,
             Zimmer,A. and Zody,M.
             Direct Submission
             Submitted (10-FEB-2000) Whitehead Institute/MIT Center for Genome
             Research, 320 Charles Street, Cambridge, MA 02141, USA
             On Mar 1, 2000 this sequence version replaced gi:6958035.
             All repeats were identified using RepeatMasker:
             Smit, A.F.A. & Green, P. (1996-1997)
             http://ftp.genome.washington.edu/RM/RepeatMasker.html
             ----- Genome Center
             Center: Whitehead Institute/ MIT Center for Genome Research
             Center code: WtBR
             Web site: http://www.seq.wi.mit.edu
             Contact: sequence_submissions@genome.wi.mit.edu
             ----- Project Information
             Center project name: L879
             Center clone name: 168_D_21
             ----- Summary Statistics
             Sequencing vector: M13; M77815; 100% of reads
             Chemistry: Dye-terminator Big Dye; 100% of reads
             Assembly program: Phrap; version 0.960731
             Consensus quality: 133287 bases at least Q40
             Consensus quality: 135023 bases at least Q30
             Consensus quality: 141009 bases at least Q20
             Insert size: 163000; agarose-fp
             Insert size: 142357; sum-of-contigs
             Quality coverage: 3.9 in Q20 bases; agarose-fp
             Quality coverage: 4.5 in Q20 bases; sum-of-contigs
             -----
             * NOTE: This is a 'working draft' sequence. It currently
             * consists of 22 contigs. The true order of the pieces
             * is not known and their order in this sequence record is
             * arbitrary. Gaps between the contigs are represented as
             * runs of N, but the exact sizes of the gaps are unknown.
             * This record will be updated with the finished sequence
             * as soon as it is available and the accession number will
             * be preserved.
             *
             * 1 2173: contig of 2173 bp in length
             * 2174 2273: gap of 100 bp
             * 2274 4319: contig of 2046 bp in length
             * 4320 4419: gap of 100 bp
             * 4420 6565: contig of 2146 bp in length
             * 6566 6665: gap of 100 bp

```

```

* 6666 9518: contig of 2853 bp in length
* 9519 9618: gap of 100 bp
* 9619 13288: contig of 3670 bp in length
* 13289 13388: gap of 100 bp
* 13389 15593: contig of 2211 bp in length
* 15600 15699: gap of 100 bp
* 15700 21304: contig of 5605 bp in length
* 21305 21404: gap of 100 bp
* 21405 26070: contig of 4666 bp in length
* 26071 26170: gap of 100 bp
* 26171 31773: contig of 5603 bp in length
* 31774 31873: gap of 100 bp
* 31874 38024: contig of 6151 bp in length
* 38025 38124: gap of 100 bp
* 38125 43487: contig of 5363 bp in length
* 43488 43587: gap of 100 bp
* 43588 49049: contig of 5462 bp in length
* 49050 49149: gap of 100 bp
* 49150 56108: contig of 6959 bp in length
* 56109 56208: gap of 100 bp
* 56209 64181: contig of 7973 bp in length
* 64182 64281: gap of 100 bp
* 64282 73216: contig of 8935 bp in length
* 73217 73316: gap of 100 bp
* 73317 81917: contig of 8601 bp in length
* 81918 82017: gap of 100 bp
* 82018 91505: contig of 9488 bp in length
* 91506 91605: gap of 100 bp
* 91606 101191: contig of 9586 bp in length
* 101192 101291: gap of 100 bp
* 101292 110389: contig of 9098 bp in length
* 110390 110489: gap of 100 bp
* 110490 120867: contig of 10378 bp in length
* 120868 120967: gap of 100 bp
* 120968 131985: contig of 11018 bp in length
* 131986 132085: gap of 100 bp
* 132086 144457: contig of 12372 bp in length.
FEATURES             Location/Qualifiers
     source            1..144457
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /map="9"
                        /clone="RP11-168D21"
                        /clone_lib="RPC1-11 Human Male BAC"
     misc_feature      1..2173
                        /note="assembly_fragment"
     misc_feature      2274..4319
                        /note="assembly_fragment"
     misc_feature      4420..6565
                        /note="assembly_fragment"
     misc_feature      6666..9518
                        /note="assembly_fragment"
     misc_feature      9619..13288
                        /note="assembly_fragment"
     misc_feature      13389..15599
                        /note="assembly_fragment"
                        clone_end:SP6
                        vector_side:right
     misc_feature      15700..21304
                        /note="assembly_fragment"
     misc_feature      21405..26070
                        /note="assembly_fragment"
     misc_feature      26171..31773
                        /note="assembly_fragment"
     misc_feature      31874..38024
                        /note="assembly_fragment"
     misc_feature      38125..43487
                        /note="assembly_fragment"
     misc_feature      43588..49049
                        /note="assembly_fragment"
     misc_feature      49150..56108
                        /note="assembly_fragment"

```

```
misc_feature 56209..64181
/note="assembly_fragment"
misc_feature 64282..73216
/note="assembly_fragment"
misc_feature 73317..81917
/note="assembly_fragment"
misc_feature 82018..91505
/note="assembly_fragment"
misc_feature 91606..101191
/note="assembly_fragment"
misc_feature 101292..110389
/note="assembly_fragment"
misc_feature 110490..120867
/note="assembly_fragment"
misc_feature 120968..131985
/note="assembly_fragment"
misc_feature 132086..144457
/note="assembly_fragment
clone_end:T7
vector_side:right"
BASE COUNT 41302 a 30843 c 30435 g 39775 t 2102 others
ORIGIN
```

```
Query Match 3.0%; Score 48.2; DB 67; Length 144457;
Best Local Similarity 57.7%; Pred. No. 0.33;
Matches 86; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1474 AGAAAAAGTGGTATCAATGTATATAAAGACAGACAGCAAGTTATGATGCAACAAATGTTTT 1533
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111287 AGTGAACCTTTGCTCTCAGAAAAAAGAAAAAGAAAAAGAAAAAGAAAAATATCACTTTT 111228

QY 1534 AGGAGATTACAATCCATGGGAGATGTATCAAACTCATCTCTATATATATATATATATCA 1593
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111227 AGAGCTTAGGATTAGGCTATATGTGAACGACACTAAATTTCTTTAAAAATATGTTTA 111168

QY 1594 ATGTTTTTAAAAAAGAAAAAAGAAAAA 1622
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111167 TTTTATTAAACAATTACAATTGGAAAAA 111139
```

Search completed: November 5, 2001, 16:38:25
Job time: 10075 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 13:52:40 ; Search time 5816.79 seconds
(without alignments)
2767.542 Million cell updates/sec

Title: US-08-894-356C-1
Perfect score: 1703
Sequence: 1 TCATTATGGAGCAATCCAA.....AAAAAAAAAAAAAAAAAAAA 1703

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	gb_est1:
2:	gb_est2:
3:	gb_est3:
4:	gb_est4:
5:	gb_est5:
6:	gb_est6:
7:	gb_est7:
8:	gb_est8:
9:	gb_est9:
10:	gb_est10:
11:	gb_est11:
12:	gb_est12:
13:	gb_est13:
14:	gb_est14:
15:	gb_est15:
16:	gb_est16:
17:	gb_est17:
18:	gb_est18:
19:	gb_est19:
20:	gb_est20:
21:	gb_est21:
22:	gb_est22:
23:	gb_est23:
24:	gb_est24:
25:	gb_est25:
26:	gb_est26:
27:	gb_est27:
28:	gb_est28:
29:	gb_est29:
30:	gb_est30:
31:	gb_est31:
32:	gb_est32:
33:	gb_est33:
34:	gb_est34:
35:	gb_est35:
36:	gb_est36:
37:	gb_est37:
38:	gb_est38:
39:	gb_est39:
40:	gb_est40:
41:	gb_est41:
42:	gb_est42:
43:	gb_est43:
44:	gb_est44:
45:	gb_est45:
46:	gb_est46:
47:	gb_est47:
48:	gb_est48:
49:	gb_est49:
50:	gb_est50:
51:	gb_est51:
52:	gb_est52:
53:	gb_est53:
54:	gb_est54:
55:	gb_est55:
56:	gb_est56:
57:	gb_est57:
58:	gb_est58:
59:	gb_est59:
60:	gb_est60:
61:	gb_est61:
62:	gb_est62:
63:	gb_est63:
64:	gb_est64:
65:	gb_est65:
66:	gb_est66:
67:	gb_est67:
68:	gb_est68:
69:	gb_est69:
70:	gb_est70:
71:	gb_est71:
72:	gb_est72:
73:	gb_est73:
74:	gb_est74:
75:	gb_est75:
76:	gb_est76:
77:	gb_est77:
78:	gb_est78:
79:	gb_est79:
80:	gb_est80:
81:	gb_est81:
82:	gb_est82:
83:	gb_est83:
84:	gb_est84:
85:	gb_est85:
86:	gb_est86:
87:	gb_est87:
88:	gb_est88:
89:	gb_est89:
90:	gb_est90:
91:	gb_est91:
92:	gb_est92:
93:	gb_est93:
94:	gb_est94:
95:	gb_est95:
96:	gb_est96:
97:	gb_est97:
98:	gb_est98:
99:	gb_est99:
100:	gb_est100:
101:	gb_est101:
102:	gb_est102:
103:	gb_est103:
104:	gb_est104:
105:	gb_est105:
106:	gb_est106:
107:	gb_est107:
108:	gb_est108:
109:	gb_est109:
110:	gb_est110:
111:	gb_est111:
112:	gb_est112:
113:	gb_est113:
114:	gb_est114:
115:	gb_est115:
116:	gb_est116:

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.


```

Qy 618 AATCTTCTTCATCTTTCGATAGATCGATAATCAAGAATCTGTATGCGCTAGAGGAACA 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 GAGCTAATTCATTTATGATAGTCCGTAGTAAGGACCCATAGGCAAGGATGTC 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 678 TTTGGACGAAGTCAAGATGTTCTTCAATGTTCTCTAGATTTGAGCAAAACCCCT 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 626 ATGTGGGAAGAAATGAGAGAAGAAATCTAAATATGCGTGACATTTAGACTCCCTCCCT 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 738 CGATTCAACAAGGTACGAGCTACATATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 797
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 686 GAA--CACAAATGTCGAGGTACATTTACTATAAACAAGATGATATGAGAACTCAAG 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 798 AACAAA 803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 744 ATTTAA 749
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
BE920468 575 bp mRNA EST 02-OCT-2000
LOCUS EST424237 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION CSTB6C8 5' sequence, mRNA sequence.
ACCESSION BE920468
VERSION BE920468.1 GI:10446544
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 575)
van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
Generation of ESTs from potato leaves and petioles
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
    source
        1..575
            /organism="Solanum tuberosum"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone="CSTB6C8"
            /clone_lib="potato leaves and petioles"
            /tissue_type="leaflets and petioles"
            /dev_stage="8 weeks old plants"
            /lab_host="SOLR"
            /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
            XhoI; Tissue was supplied by Dr. Fry (Cornell University).
            Leaflets and petioles were isolated from 8 week old
            greenhouse grown plants. The plants were watered and
            fertilized freely. The tissue was immediately frozen in
            liquid nitrogen."
BASE COUNT 167 a 77 c 139 g 192 t
ORIGIN
    1..575
        /organism="Solanum tuberosum"
        /cultivar="Kennebec"
        /db_xref="taxon:4113"
        /clone="CSTB6C8"
        /clone_lib="potato leaves and petioles"
        /tissue_type="leaflets and petioles"
        /dev_stage="8 weeks old plants"
        /lab_host="SOLR"
        /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
        XhoI; Tissue was supplied by Dr. Fry (Cornell University).
        Leaflets and petioles were isolated from 8 week old
        greenhouse grown plants. The plants were watered and
        fertilized freely. The tissue was immediately frozen in
        liquid nitrogen."
Query Match 8.3%; Score 141; DB 142; Length 575;
Best Local Similarity 58.4%; Pred. No. 9.2e-27;
Matches 331; Conservative 0; Mismatches 215; Indels 21; Gaps 4;

Qy 833 AATACGTGTACAGACGTTCAACATGACGTGTGGATAGTATGATGACATGCGTCAAAATC 892
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6 AACTCATGTAAACATCTTTTACTGTAACGAGTGCTTATGTATGGACTTGCCTGTGATAAAATC 65
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 893 AAAAGATGACCTCGTATCAGAGGATCATCGAACGCAAAATCAGCTCGAGTACTCAG 952
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 AAAGACCGGATTGG---AGAGAAGAAATCATAGATGATAGTATGATGAGTCTTTGG 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 953 TTTTACAGCGGATTGCCGAGACTTCTGACGCCGCCCGCTGTCGCCCTAACTACTTTGGCAA 1012
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 123 ATGTGCTGGAGATTTTAGAGCAGCATTCATCAATCCACCACCTTCCTCAATCTTATTTGGAA 182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1013 CTGTCTTTCGCTCATGCGCTTCAAAAGCAACA--CATAAAGAGTTAGTTGGGATAAAG 1069
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 TTGTGTAGTTGGGTGTGTCACAAGATCAATTAAGGCATGTTGACTTAGTTGGAAAGGAGG 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1070 GCTTCTTGTGTCAGTTGCGAGCTATTGAGAGGCCATTGAAAAGAGGTTGCACACAGAAA 1129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 TTTTGAATTCGCGTGAATTAATTGAGAAGTCAATTCAAAAAATAAGAGATGAAGA 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1130 AGGCGTTCTTTCGAGATGCAAAACTTGGTTATCGGAATCTAATGGAATCCCTTCAAAAG 1189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 ATGGTCTCTTAATGGTG-----ATTGGTTAAAAGAGTTGGACAACGTAGACCTTAATTCG 356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1190 ATTTCTCGGATTACCGGATCGCCTAAGTTCGATTCGATTCGATTCGATTCGATTCGATTCG 1249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 357 ATTTTCTCAATTCGTCGATCGCCAAACATGACTTCTATGCTGCTGATTTTGGATGGG 416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1250 AAGCCTGCAAAATTTGACATTTACCTCTGTTGATTATGACGAA-----TTGATTTA 1300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 417 AAGGCCGTGAAGTTAGAAATTCATTTCTTGACAATGATGATGGTGAATTCGATGTC 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1301 TGTGATTCAAGTCCAGGATTTTGAAGAGGTGTGGAGATTTGGAGTATCATTTGCCCTAAGAT 1360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 477 TCTTAGTAAATCTAAGAAATTTTGATGGAGATTTAGAGATTTGGTTGCTTTGCTATTATAC 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1361 TCATATGGATGCATTTTGCACAAATCTTT 1387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 537 TCAATGAATGCTTTTTCGTCGTATATT 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
BG599447 690 bp mRNA EST 12-APR-2001
LOCUS EST504342 cSTS Solanum tuberosum cDNA clone cSTS25I22 5' sequence,
DEFINITION mRNA sequence.
ACCESSION BG599447
VERSION BG599447.1 GI:13616583
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 690)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.
FEATURES
    source
        1..690
            /organism="Solanum tuberosum"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone="cSTS25I22"
            /clone_lib="cSTS"
            /tissue_type="sprouting eyes from tubers"
            /dev_stage="12-14 weeks post harvest"
            /lab_host="SOLR"
            /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
            XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
            taken from tubers. The tubers were incubated at 26C in the
            dark for 2-3 weeks prior to sprouting. The eyes were
            frozen in liquid nitrogen immediately upon removal from
            tubers."
BASE COUNT 178 a 156 c 144 g 212 t
ORIGIN
    1..690
        /organism="Solanum tuberosum"
        /cultivar="Kennebec"
        /db_xref="taxon:4113"
        /clone="cSTS25I22"
        /clone_lib="cSTS"
        /tissue_type="sprouting eyes from tubers"
        /dev_stage="12-14 weeks post harvest"
        /lab_host="SOLR"
        /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
        XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
        taken from tubers. The tubers were incubated at 26C in the
        dark for 2-3 weeks prior to sprouting. The eyes were
        frozen in liquid nitrogen immediately upon removal from
        tubers."

```

Query Match	7.5%;	Score 127;	DB 155;	Length 690;	
Best Local Similarity	52.9%;	Pred. No. 5.1e-23;			
Matches 369;	Conservative 0;	Mismatches 315;	Indels 13;	Gaps 4;	
QY 40	AATGCCAAGTTACACACCATCTGACACAAAGAGTGTCCAGTTATCGCTACCGGTAACAT	99			
Db					
4	AATGTCAGTTGCGCCACCTCCGGGGGGCA--CGGAGGTGACATCCCTCTTACTT	60			
QY 100	TCTTCGATATCCCTGGTTGCACTTGAATPAAGATGAGCCCTCTCTGTTTTACGACTTTC	159			
Db					
61	ATTTTGCATGTTGGTTAGGTTAGGTTCCACCGTATACGGGGATATATTCTACAAGCTCC	120			
QY 160	CGTACCCAGAACACATTTCTTGGACACTGTTATCCCTAATCTTAAGGCTCTTTGTCTC	219			
Db					
121	CAATTTCCAAACCGGATTTGTTTCAAAACATTTATCTCTCTCTTAAAAATTCACCTCC	180			
QY 220	TCACCTCTAAACACACTACGTTTCGGCTTAGCGGAATTTGTTGATGCCGATCAAAATCGGCG	279			
Db					
181	TCACCTCTCAACACATACGCCCTTAGCTGGAACGTTCTGCTCCACTAGATACAAACG	240			
QY 280	AAATGCCGAGTTTCAGTACTCCGTTGATGAGGGGACACTGCATTAATTTGATCGTTGCGG	339			
Db					
241	GATATCTCTGAGTTACATTA-----TGTGACAGAGATCTCTATCTGTTACTTTTTCTG	294			
QY 340	AGTCTGACCAGGATTTTGACTACCTTAAGGTCATCAACTGGTAGATTTCCAATGTTTGC	399			
Db					
295	AGACTGATATGGATTTCAATATATCATTTGTTGAGACCATCCGGGTAAATGCTAAGGATTTT	354			
QY 400	ATGGCCTTTTATGTTATGCCACGGCTTATAAGGACCATGCAAGACTATAAAGTAGTCC	459			
Db					
355	AT--CACTTTGTCTACGTTTAGGGACCTAAGGATCACCGGGTCCAAATTAGCTC	411			
QY 460	CGCTGACCGGTGCAAGTAACGGTTTTTCCCTAACCGTGGCATAGCCGTGCTCGACGG	519			
Db					
412	CGGTCTTAGCCATCAAGTGACACTTTTTCCGAATCTTGGTGTATCCATTTGTTTCACTA	471			
QY 520	CACATCAATTCATTCGAGATGCTAAAGTTTTGTAATGTTTCATCAATGCTTGGGCTATA	579			
Db					
472	ACATCATGTTGTTGGTATGAGCTACCATAGTAGGGTTTCAATAGGGGCTGGGCTCTAC	531			
QY 580	TTAACAAATTTGGGAAGACGGGACTTCTTGTCGCGAATCTCTTCCATCTTTTCGATA	639			
Db					
532	TCCATAAATTCGGTGGAGATGAACAATCTTATCGAATGAGCTAATTCATTTATGATA	591			
QY 640	-GATCATTAATCAAGATCTGATGCCCTAGAGGAACAATTTTGGAACGAATGCAAGAT	698			
Db					
592	GGTCCGTAGTAAAGACCCATATGGACAAGGATGTTTCATATGGGAAGAAATGAAGAA	651			
QY 699	GTTCTTGAATGTTCTCTAGATTGGGAAGCAAAACCC 735				
Db					
652	AGATCAAGATATGCGTGACATATGACTCTCTCTCC 688				
RESULT 4					
BE434257					
LOCUS	BE434257 591 bp mRNA EST 24-JUL-2000				
DEFINITION	EST405335 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA clone cLEG15H6, mRNA sequence.				
ACCESSION	BE434257				
VERSION	BE434257.1 GI:9432100				
KEYWORDS	EST.				
SOURCE	tomato.				
ORGANISM	Lycopersicon esculentum				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	Asteridae; euasterids I; Solanales; Solanaceae; Solanum;				
	Lycopersicon.				
REFERENCE	1 (bases 1 to 591)				
AUTHORS	Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,				
	Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,				
	Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley				

'S.D.	Generation of ESTs from tomato fruit tissue, breaker stage				
Unpublished (2000)					
Contact: David Frisch					
Clemson University Genomics Institute					
Clemson University					
100 Jordan Hall, Clemson, SC 29634, USA					
Tel: 864 656 4366					
Fax: 864 656 4293					
Email: dfrisch@clemson.edu					
5 prime sequence.					
FEATURES	Location/Qualifiers				
1..591					
/organism="Lycopersicon esculentum"					
/cultivar="TA496"					
/db_xref="taxon:4081"					
/clone_lib="cLEG15H6"					
/clone_lib="tomato breaker fruit, TIGR"					
/tissue_type="pericarp"					
/dev_stage="breaker"					
/lab_host="SOLR"					
/note="Vector: pBluescriptSKmCudapt; Site_1: EcoRI;					
Site_2: XhoI; Fruit were harvested at the breaker stage					
(first sign of lycopene accumulation on the blossom end					
of the fruit). Fruit were cut in half and the seeds and					
locules were discarded prior to freezing the pericarp."					
BASE COUNT	159 a 129 c 121 g 182 t				
ORIGIN					
Query Match	7.2%;	Score 122.8;	DB 167;	Length 591;	
Best Local Similarity	53.7%;	Pred. No. 6.5e-22;			
Matches 303;	Conservative 0;	Mismatches 252;	Indels 9;	Gaps 2;	
QY 132	ATGCAGTCCCTCTCTGTTTACGACTTTCGGTACCCAAAGACACATTTCTTGACACTGTT	191			
Db					
7	ATGAGCGGATATATTATTTACAAGCTCTCCATTTTCAAAACCCGATTTCTGTTCAAAACAT	66			
QY 192	ATCCCTAATCTTAAGGCTCTTTGCTCTCACTCTAAACACTACGTTCCGCTTAGCCGA	251			
Db					
67	ATTCTCTCTCTTAAAAATTCACCTCTCCCTCACTCTCAAAACACTATACGCCCTTAGCCGA	126			
QY 252	AATTTCTTGATGCCGATCAATCGGGGAAATGCCGAAGTTTCAGTACTCCGCTGATGAG	311			
Db					
127	ACGTTGCTTGTCCACTAGTACAAACGGATATCCCTGAGTTAGCTTA-----TGTGACA	180			
QY 312	GGGACTCGATTAATTTGATTCGTTGGGAGTCTGACGAGATTTTGACTTAAAGGT	371			
Db					
181	GGAGATCTCTGCTCTCTTCTTTTCGAGACTGATGAATTTCAATTTATCTCATTTGTT	240			
QY 372	CATCAACTGCTAGATTCGAATGATTTGCGCTTTTATGTTATGCGCGGTTATA	431			
Db					
241	GACCATCCCGGTAAAGCTAAGGATTTTAT---CACTTTGTTCTAAGTTAGGGAACT	297			
QY 432	AGGACCATCAAGACTATAAAGTGTATCCCGCTGTCAGCGTGCAGTAACCGTTTTCTCT	491			
Db					
298	AAGGATGCACCGGGTCCCACTAGCCCGCTCTTAGCCATTCAGCTGACACTTTTCCG	357			
QY 492	AACCGTGGATAGCCCTGGCTCTGACGGGACATCATTAATTCAGATGCTTAAAGTTT	551			
Db					
358	AATCTTGGTGTATCCCATTTGTTTCACTAACCACTCATGTTGTTGGTGTAGGACTATA	417			
QY 552	GTAAATGTTTCATCACTTGGGCTATATTAACAATTTGGGAAACACCGGACTTGTG	611			
Db					
418	GCAGGTTTCATTAAAGCGGTGGGCTCTACTCCCAAAATTCGGTGGACATGAACAATTTCTTA	477			
QY 612	TCCGCAATCTTCTTCCATCTTCGATAGATCGATTAATCAAGATCTGATGCCCTTAG	671			
Db					
478	TCAATGACCTAATTCATTTTATGATAGTCCGTAGTAAAGACCCATATGGCAAGGG	537			
QY 672	GAACATTTTGGAAACGAATGCAA 695				
Db					
538	ATGTCATCTGGGAAGAAATGAAA 561				

QY	490	CTAACCGTGGCATAGCCGTGGCTCTGACGGGCACATCAATCAATTCAGATGCTAAAGTT	549
Db	359	CGAATCTTGGCATATCCGTTGGTTTTAGTAACCATCATGTCGCTGTGTGAATAACCA	418
QY	550	TTGTAATGTTTCATCAATGCTTGGGCTATATTAACAAATTTGGGAAGACCGGACATTGT	609
Db	419	TCGTGAAATTCATAGAAGATGGGTTTACTCAACAAATTCGGCGGTGATGAACAGTCT	478
QY	610	TGTCCCGGAATCTTCTTCATCTTCGATAGATCGATAATCAAGATCTGTATGGCCTAG	669
Db	479	TAGAGATGATGTTTCATCTTCATCTTCGATAGATCGATAATCAAGATCTGTATGGCCTAG	538
QY	670	AGGAAACATTTTGGACGAATGCA	694
Db	539	GGACGATTATATGGGATGAATGAA	563
RESULT	6		
LOCUS	BE921494		
DEFINITION	EST425179 potato leaves and petioles Solanum tuberosum cDNA clone		
ACCESSION	BE921494		
VERSION	BE921494.1	GI:10447486	
KEYWORDS	potato.		
SOURCE	Solanum tuberosum		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.		
REFERENCE	1 (bases 1 to 564)		
AUTHORS	van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J., Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.		
TITLE	Generation of ESTs from potato leaves and petioles		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Cathy Ronning The Institute for Genomic Research For clone request: please contact Research Genetics, Libraries Division tel 1-800-711-6195, email cdna@resgen.com.		
FEATURES	Location/Qualifiers		
source	1..564		
	/organism="Solanum tuberosum"		
	/cultivar="Kennebec"		
	/db_xref="taxon:4113"		
	/clone="CSTB13B18"		
	/clone_lib="potato leaves and petioles"		
	/tissue_type="leaflets and petioles"		
	/dev_stage="8 weeks old plants"		
	/lab_host="SOLR"		
	/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Tissue was supplied by Dr. Fry (Cornell University). Leaflets and petioles were isolated from 8 week old greenhouse grown plants. The plants were watered and fertilized freely. The tissue was immediately frozen in liquid nitrogen."		
BASE COUNT	150 a 139 c 97 g 178 t		
ORIGIN			
Query Match	7.0%;	Score 119.6;	DB 142; Length 564;
Best Local Similarity	53.6%;	Pred. No. 1.9e-21;	
Matches	303; Conservative	0; Mismatches 250;	Indels 12; Gaps 2;
QY	130	AGATGCGAGTCCCTCTGTTTACGACTTCCGTCACCAAGAACACATCTTCTTGGACACTG	189
Db	11	ATATGGTCGGATTTATCTACAGCTACCAATTTCCAACTCATGTTCTGTCNAACCA	70
QY	190	TTATCCCTAATCTAAGGCTCTTTGTCCTCTCACTCTAAACACACTACGCTTCGGCTTAGCG	249
Db	71	TTATTCTCTACTTAAACATCTACTCTCCCTCACTCTCAACACTACACACCTTAGCTG	130
QY	250	GAAATTTGTTGATGCGCATCAATCGGGCAATGCGGAGTTTCAGTACTCCGCTGATG	309
Db	131	CGAACATGCTTCTCCACTAAATTCAGTGGTATCTCTGAGTTGCGGTTA-----TGTTGA	184
QY	310	AGGCGACTCGATAACTTTGATGCTTGGCGAGCTCTGACCGAGATTTTCAGTACCTTAAG	369
Db	185	CTGGAGATCTATATCTGTTACTTTTATTGACACTGATGATGATTTCAATCATCTCATG	244
QY	370	GTATCAACTGGTAGATTCATGATTTTCATGCGGCTTTTTTATGTTATGCCACGGGTTA	429
Db	245	GTAACATCTCGAAATGCTAAGGATTTTATCCCATCTCCTCAATTTGGCACAACC-----	299
QY	430	TAGGACCATGCAAGACTATAAGTATGATCCCGCTCGTAGCCGTGCAAGTAACCGTTTTC	489
Db	300	-TAAGGATGCACCGGGGGTCAACACTAGTCCCGGCTCTTAGGCCATTCAGTGACACTTTTC	358

Db 143 AGCTACCAATTTCCAAACTCGATTTTCGTTCAAAACCATTTATTCCTACTCTTAACACATTTAC 202
QY 214 TGCTCTCTCACTCTAAACACTAGCTTCCGCTTAGCCGGAATTTGTTGATGCCGATCAAT 273
Db 203 TCTCCCTCACTCTCAAAACACTACACACCTTAGCTGGCAACATTTGCTTGCCACTAAAT 262
QY 274 CGGGGAAATGGCGAAGTTTCAGTACTCCGCTGATGAGGCGGACTCGATAAATTTGATCG 333
Db 263 CGAGTGGTTATCTGAGTTGCGTTA-----TGTGACTGGAGATTCATATCTGTTACTT 316
QY 334 TTGGGAGTCTGACAGGATTTTGACTACTCTTAAGGTCATCAACTGGTAGATTCOAATG 393
Db 317 TTATTGAGACTGATATGAATTTCAATCATCTCATTTGGTAACCATCTCTCGAAATGCTAAGG 376
QY 394 ATTTGATGCGCTTTTATGTTATCCACGGGTTATAGGACCATGCAAGACTATTAAG 453
Db 377 ATTTTATCCCATCTCTCAATTTGGCAAC-----CTAAGGATGACCCGGGGTCAAC 430
QY 454 TGATCCCGCTCGTAGCGTGCAAGTAACCGTTTTTCTTAACCGTGATGCGTGGCTC 513
Db 431 TAGTCCCGCTTAGCCATTCAGTGACACTTTTCCGGAATCTTGGCATATCCGTTGGTT 490
QY 514 TGACGGACATCATTCATTTGCGAGATGCTTAAAGTTTGTATGTTTCATCAATGCTTGGG 573
Db 491 TTAGTAACCATCATCTGCTGTTGTGGTGAAGTAACCATGCTGGAATTCATAAGAACAATGGG 550
QY 574 CCTATATTACAAA 587
Db 551 GTTTACTCAACAAA 564

RESULT 7
BC097054 508 bp mRNA EST 29-JAN-2001
LOCUS EST461573 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION cSTB45P9 5' sequence, mRNA sequence.
ACCESSION BC097054
VERSION BC097054.1 GI:12587089
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
REFERENCE
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel.1-800-711-6195, email cdna@resgen.com.

FEATURES
source
1. .508
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB45P9"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site.1: EcoRI; Site.2:
XhoI. Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 127 a 130 c 87 g 164 t
ORIGIN

Query Match 6.5%; Score 111.4; DB 173; Length 508;
Best Local Similarity 54.0%; Pred. NO. 6.8e-19; Mismatches 226; Indels 9; Gaps 2;
Matches 276; Conservative 0;
QY 36 GAAAAATGCAAGTTACACACCATCTGACACAAAGATGTCGAGTATATGCGTACCGGTA 95
Db 2 GAGCAATGCCAAGTTGCGSCACCTCCACGGGCGCAACAGAGCTCCCTCTA 61
QY 96 ACATTTCTCGATATCCCTGGTTGCACTTGAATAAGATGAGTCCCTCTGTTTATAGGAC 155
Db 62 ACTATTATTTGATCATATATGTTTGGCTATATGCTGCGGATTTTATTTCTACAAG 121
QY 156 TTTCCGTAACCCCAAGAACACATTTCTTTGGACACTGTTATCCCTAATCTTAAAGGCTCTTTG 215
Db 122 CTACCAATTTTCCAACTCGATTTTCGTTTCAACCATTTATTCCTACTCTTAAACATTCAC 181
QY 216 TCTCTCACTCTTAAACACTACGTTCCGCTTAGCGGAAATTTGTTGATGCCGATCAATCG 275
Db 182 TCCTCACTCTCAACACTACACACCTTAGCTGGCAACATTCCTTGTCCACTAAATCG 241
QY 276 GCGGAAATGCCGAAGTTTCAGTACTCCGCTGATGAGGGGACTCGGATACTTTGATCGTT 335
Db 242 AGTGGTTATCTCTGAGTTGCGTTA-----TGTGACTGGAGATTCATATCTGTACTTTT 295
QY 336 GCGGAGTCTGACACGAGATTTTGACTACCTTAAAGGTCATCAACTGGTAGATTTCAATGAT 395
Db 296 ATTGAGACTGATATGAATTTCAATCATCTCAATTTGGTAACCATCTCTCGAAATGCTAAGGAT 355
QY 396 TTGCATGGCCCTTTTATGTCGACACGGGTTATAGGACCATGCAAGACTATAAAGTG 455
Db 356 TTTTATCCCTTCATTCCTCAATTTGGCACACCTAAGGATGCACCGGGGTC---AAACTA 412
QY 456 ATCCCGCTGCTAGCCGTGAAGTAACCGTTTTTCTTAACCGTGGCATAGCCGTGGCTCTG 515
Db 413 GTCCCGCTCTTAGCCATTCGAAGTGACACTTTTTTCCGAAATCTTGGCATATCCGTTGGTTT 472
QY 516 ACGGCACATCATTCATTCAGATGCTAATAA 546
Db 473 AGTAACCATCATCTCGCTTGTGATGAATA 503

RESULT 8
BF051101 578 bp mRNA EST 16-OCT-2000
LOCUS EST436276 tomato developing/immature green fruit Lycopersicon
DEFINITION esculentum cDNA clone cLEM21H7 5' sequence, mRNA sequence.
ACCESSION BF051101
VERSION BF051101.1 GI:10804997
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
REFERENCE
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
TITLE Generation of ESTs from tomato fruit tissue, immature green
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu.
FEATURES
source
1. .578
Location/Qualifiers

/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEM21H7"
/clone_lib="tomato developing/immature green fruit"
/tissue_type="fruit"
/dev_stage="immature green (5-35 days post-anthesis)"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmCUadapt; Site_1: EcoRI;
Site_2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and
harvested at 7 day intervals through 35 dpa. Equal masses
of tissue from each stage were combined (including seeds
and locules) prior to mRNA isolation."
BASE COUNT 192 a 88 c 135 g 163 t
ORIGIN

Query Match 6.5%; Score 111.4; DB 143; Length 578;
Best Local Similarity 54.7%; Pred. No. 7e-19;
Matches 317; Conservative 0; Mismatches 241; Indels 21; Gaps 4;
Qy 546 AGTTTGTGAATGTCATCAATGCTTGGCCCTATATTAACAATTTGGGAAGACGCGAC 605
Db 5 ACTATAGCAGGGTTCATTAAGCGGTGGGCTCTACTCCAAATTCGGTGGACATGAACAA 64
Qy 606 TTGTTGTCGGGAATCTTCTTCCATCTTTTCGATAGATCGATAATCAAAAGATCTGTATGGC 665
Db 65 TTCATTATGATGAGTAAATTCATTTATGATAGTCCGTAGTAAAGACCCATATGA 124
Qy 666 CTAGAGGAACATTTTGGGAAGAAATCAAGATGTTCTTTGAAATGTTCTCTAGATTTGGA 725
Db 125 CAAGGATGTCATCTGGGAAGAATGAAAGCAAGATCTAGATATGCGTGACGTT--- 181
Qy 726 AGCAAAACCCCTCGATTCAACAAGTAGAGCTACATATGCTCTCCCTTGCTGAATC 785
Db 182 ATGACTCTCTCTGGAACAAGGTCGAGGTACATTTACTATTAAGAAGATGAATA 241
Qy 786 CAGAGATGAAGAACAAAGTACTGAATCTCAGAGGATCCG-----AACCGACA 833
Db 242 GAGAACTGAAGAAATTTATATGAACCTCAAGAGACGGGTAGTAGTACTACTCTA 301
Qy 834 ATACGTGTAACGAGCTTCACAAATGACGTGGATACGTATGATGACATGATGCTCAATCA 893
Db 302 ACTCATGTAACATCTTTTACTGTAAACGAGTCTTATATATGAGCTTGTGTAATAATC- 360
Qy 894 AAAGATGACGTCTATCAGAGGAATCATCGAAGCAAGAAATGAGCTCGAGTACTTCACT 953
Db 361 --AGAGGACCGATAGGAGAAGATGATCATAGATGATAGTGAATGTTTCGGA 418
Qy 954 TTTACAGCGGATTCGCGAGGACTTCTGACGCCCGCGTTCGCCCTAACTACTTTGGCAAC 1013
Db 419 TGTGACGAGATTTTAGAGCGGATTTCAATCCACACTTCTCTCAATCTTATTTGGGAAT 478
Qy 1014 TGTCTTGGCTATCGGTGCAAAAGCAACA---CATAAAGAGTTAGTTGGGGATAAAGGG 1070
Db 479 TGCATAGTTGGGTATGTCACAAAATCAATAAGGCTGTGACTAGTTAGTGAAGGGAAGGG 538
Qy 1071 CTCTCTTGTTCAGTTGACGCTATTGGAGAAGCCATTGAA 1109
Db 539 TTTAAATTCGGTAGATCAATTTGGAGAAGTCAATCAA 577

RESULT 9
LOCUS AW221049 606 bp mRNA EST 07-DEC-1999
DEFINITION EST297518 tomato fruit mature green, TAMU Lycopersicon esculentum
cDNA clone cLEF3f1, mRNA sequence.
ACCESSION AW221049
VERSION AW221049.1 GI:6532733
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 606)
Alcala, J., Vrebalov, J., White, R., Matern, A.L., Holt, I.E., Liang, F.,
Upton, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Ronning, C.M.,
Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.

FEATURES
source
1..606
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF3f1"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
BASE COUNT 147 a 143 c 123 g 193 t
ORIGIN

Query Match 6.4%; Score 109.8; DB 113; Length 606;
Best Local Similarity 52.2%; Pred. No. 1.9e-18;
Matches 321; Conservative 0; Mismatches 282; Indels 12; Gaps 3;
Qy 19 AAATGGTGAAGTCTTTGAAAAATGCCAAGTTACACCACCATCTGACACACAGATGCG 78
Db 2 ACATGGCCACCGTATTGAGCAATGTCAAGTTGGCCACCTCCCGGGCGCAACGGAGG 61
Qy 79 AGTTATCGCTACCGGTAACTTCTCGATATCCCTGCTGCACCTTGAATGAATGCACT 138
Db 62 TGATA---CTCCCTCTACTTATTTTGACCATGTTGTTAGGTTTCGCCGTATGAGGC 118
Qy 139 CCCTTCTGTTTACGACTTTCCTGACCAAGAACACATTTCTTGACACTGTTATCCCTA 198
Db 119 GGATATTATTTTACAAGCTCTCCATTTTCAAAACCGGATTTCTGTTCAAAACATATTCTC 178
Qy 199 ATCTTAAGCCCTCTTTGCTCTCACTCTAAACACTAGTTCCCGTTAGCGGAATTTGT 258
Db 179 CTCTTAAAAATTCATCTCCCTCACTCTCAACACACTACACGCTTAGCCGGAACGTTG 238
Qy 259 TGATGCGATCAATTCGGGCGAAATGCGAAGTTTTCAGTACTCCGCTGATGAGGCGACT 318
Db 239 CTTGCTCACTAGATACAAACGGATATCTTGAGTTACGTTA-----TGTGACAGAGATT 292
Qy 319 CGATAAATTTGATCGTTGCGGAGTCTGACCAGGATTTTGACTACCTTAAAGGTCAATCAAC 378
Db 293 CTGTGCTGTGTACTTTTTCGAGACTGATATGAATTTCAATTCATCTGATGACCATC 352
Qy 379 TGGTAGATTCGAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 438
Db 353 CGCGTAAGGCTAAGGATTTTAT---CACTTTGTTCTTAAGTTAGGGGAACCTAAGGATG 409
Qy 439 TGCAAGACTATAAAGTATGATCCCGCTGAGCGGTGCAAGTAACCGTTTTCCTAACCGTG 498
Db 410 CACCGGGGTCCAACCTAGCCCGCTCTTAGCCATTGAGTGACACTTTTTCGGAATCTTG 469

```

QY 499 GCATAGCCGGTCTGCTGACGGCACATCAATTCAATTCAGATGCTAAAGTTTGTAAATGT 558
    || || || || || || || || || || || || || || || || || || || || || ||
Db 470 GTGTATCCATTTGTTTCCATTAACCATCTGTTGTGTGATGAGCTACTATAGCAGGTT 529
    || || || || || || || || || || || || || || || || || || || || || ||
QY 559 TCATCAATGCTTGGGCTATATTAAACAAATTTGGGAAGACGCGGACTTGTTCGCCGA 618
    || || || || || || || || || || || || || || || || || || || || || ||
Db 530 TCATTAAAGGCTGGGCTCTACTCCACAAATTCGGTGGACATGAACAATCTTTATCGAATG 589
    || || || || || || || || || || || || || || || || || || || || || ||
QY 619 ATCTTCTTCCATCTT 633
    || || || || || || || || || || || || || || || || || || || || || ||
Db 590 AGCTAATCCATTTT 604
    || || || || || || || || || || || || || || || || || || || || || ||

RESULT 10
LOCUS BE922784
DEFINITION EST426553 potato leaves and petioles Solanum tuberosum cDNA clone
        CSTB22E2 5' sequence, mRNA sequence.
ACCESSION BE922784
VERSION BE922784.1
KEYWORDS GI:10448860
SOURCE EST.
ORGANISM Solanum tuberosum
REFERENCE 1. .570
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
        Uterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
        ,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
        The Institute for Genomic Research
        For clone request: please contact Research Genetics, Libraries
        Division tel 1-800-711-6195, email cda@resgen.com.

FEATURES
    source
        1. .570
            /organism="Solanum tuberosum"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone="CSTB22E2"
            /clone_lib="potato leaves and petioles"
            /tissue_type="leaflets and petioles"
            /dev_stage="8 weeks old plants"
            /lab_host="SOLR"
            /note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
            XhoI; Tissue was supplied by Dr. Fry (Cornell University).
            Leaflets and petioles were isolated from 8 week old
            greenhouse grown plants. The plants were watered and
            fertilized freely. The tissue was immediately frozen in
            liquid nitrogen."
BASE COUNT 139 a 136 c 115 g 180 t
ORIGIN

Query Match 6.3%; Score 107.8; DB 142; Length 570;
Best Local Similarity 53.1%; Pred. No. 6.3e-18;
Matches 304; Conservative 0; Mismatches 257; Indels 12; Gaps 3;

QY 2 CATTATGAGCAATCCAAATGGTGAAGGTTCTTTGAAAAATGCCAAGTTACACACCATC 61
    || || || || || || || || || || || || || || || || || || || || || ||
Db 5 CATTATAGTACTAAATAACATGCGTCTTTGATAGAGCAATGTCAGTTCCGCCACCTCC 64
    || || || || || || || || || || || || || || || || || || || || || ||
QY 62 TGACACACAGATGTCGAGTTATCGCTACCGGTAAACATTTCTTCGATATCCCTGGTTGCA 121
    || || || || || || || || || || || || || || || || || || || || || ||
Db 65 GGGCGGCGCA---GCGAGGTGACACTCCCTCTACTATTATTTGATCATGTTTGGTTAGG 121
    || || || || || || || || || || || || || || || || || || || || || ||
QY 122 CTTGAATAAGATGACAGTCCCTCTGTGTTTACGACTTTCCGTACCAAGAACACATTTCTT 181
    || || || || || || || || || || || || || || || || || || || || || ||
Db 122 GTTCCACCGTATAGCGCGGATATTATTCTAAAGTCCCTCCCATTTCCAAACCGGATTTCT 181
    || || || || || || || || || || || || || || || || || || || || || ||
QY 182 GGACACTGTTATCCCTAATCTTAAGGCTCTTTGTCTCTCACTCTAAACACTACGTTCC 241
    || || || || || || || || || || || || || || || || || || || || || ||

```

```

Db 182 TCAAAACATTATTCTCTCTTAAATAATTCACCTCTCCTCACTCTCAACACTATAGCC 241
    || || || || || || || || || || || || || || || || || || || || || ||
QY 242 GCTTAGCGGAAATTTGTTGATCGCGATCAATCGGGGAAATGCCGAAGTTTTCAGTACTC 301
    || || || || || || || || || || || || || || || || || || || || || ||
Db 242 CTTAGCTGGAACGTTGCTCGTCCACTAGATACAAACGGATATCTCGAGTTACATTA--- 298
    || || || || || || || || || || || || || || || || || || || || || ||
QY 302 CCGTGTAGAGGGGACTCGATAACTTTTGATCGTTGCGGAGTCTGACCAGGATTTTGACTA 361
    || || || || || || || || || || || || || || || || || || || || || ||
Db 299 ---TGTACAGGAGATTCCTATCTGTTACTTTTCTGAGACTGATATGATTTCAATTA 355
    || || || || || || || || || || || || || || || || || || || || || ||
QY 362 CTTAAAGGTCATCAACIGGTAGATCCCAATGATTTTCATGGCCCTTTTATGTTATGCC 421
    || || || || || || || || || || || || || || || || || || || || || ||
Db 356 TCTCATTTGGTGACCATCCGCGTAATGCTAAAGGATTTTAT---CACTTTGTCTCAGGTT 412
    || || || || || || || || || || || || || || || || || || || || || ||
QY 422 ACCGGTTATAAGGACCATGCAAGACTATAAAGTCATCCCGCTCTAGCCGTCGAAGTAAC 481
    || || || || || || || || || || || || || || || || || || || || || ||
Db 413 AGGGGAACCTTAAGGATGACACCCGGGGTCCAATTAGTCCGGTCTTAGCCCATTCAGTGAC 472
    || || || || || || || || || || || || || || || || || || || || || ||
QY 482 CGTTTTCCTTAACGCTGGCATAGCCGTGCTCTGACGGCACATCATTTCAATTGCGAGATGC 541
    || || || || || || || || || || || || || || || || || || || || || ||
Db 473 ACTTTTCCGAATCTTGGTGTATCCATTGGTTTCACTAACCATCATCTGTTGTTGATGG 532
    || || || || || || || || || || || || || || || || || || || || || ||
QY 542 TAAAGTTTGTAAATGTTCAATGCTTGGGC 574
    || || || || || || || || || || || || || || || || || || || || || ||
Db 533 AGCTACCATAGTAGGTTTCATTAGGCGCTGGGC 565
    || || || || || || || || || || || || || || || || || || || || || ||

RESULT 11
LOCUS AW616119/c
DEFINITION hirsutum trichome, Cornell University Lycopersicon
ACCESSION AW616119
VERSION AW616119.1
KEYWORDS GI:7322152
SOURCE EST.
ORGANISM Lycopersicon hirsutum.
REFERENCE 1 (bases 1 to 603)
AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang
        F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W.,
        Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
        Generation of ESTs from wild tomato (Lycopersicon hirsutum)
        trichomes
        Unpublished (2000)
        Contact: David Frisch
        Clemson University Genomics Institute
        Clemson University,
        100 Jordan Hall, Clemson, SC 29634, USA
        Tel: 864 656 4366
        Fax: 864 656 4293
        Email: dfrisch@clemson.edu
        3 prime sequence.
FEATURES
    source
        1. .603
            /organism="Lycopersicon hirsutum"
            /db_xref="taxon:62890"
            /clone="cLHT6C7"
            /clone_lib="L. hirsutum trichome, Cornell University"
            /tissue_type="trichome"
            /dev_stage="mixed stages"
            /note="Leaves of various stages were shaken in liquid
            nitrogen, shearing off trichomes. This procedure yielded a
            mixture of cells which is highly enriched for trichome
            likely with minor contaminations of other types of leaf
            cells."
BASE COUNT 209 a 122 c 90 g 182 t
ORIGIN

```

```
Query Match      6.3%; Score 107.6; DB 118; Length 603;
Best Local Similarity 57.2%; Pred. No. 7.3e-18;
Matches 263; Conservative 0; Mismatches 179; Indels 18; Gaps 3;

Qy 963 GATTGCCGAGACTTCTGACGCCCGGTCGCGCTACTACTTTGGCAACTGCTTGGC 1022
Db 602 GATTTCAGGCGGATTCATACCCACCTCTCTCAATCTTATTTGGGAATGTCATAGT 543
Qy 1023 TCATCCGTTGCAAAAGCACACA--CATAAAGACTAGTTGGGGATAAAGGCTCTTGTT 1079
Db 542 GGGTATGTCACAAAATCAACAAGCGATGTTGACTTAGTTGGAAAGGAGGCTTTAAAAAT 483
Qy 1080 GCAGTTGCAGCTATTGGAGAGACCATCAAAAAGAGTTGCAACAACGAAAAAGCGCTCTT 1139
Db 482 CGGTAGAAATCAATGGAGAGTCAATCAAGAAAAATGAAGGATGATGATGGTCTT 423
Qy 1140 GCAGATGCAAAACTTGTTTATCGGAATCTTAATGGAATCCCTTCAAAAAGATTTCTCGG 1199
Db 422 AATGGTG-----ATTGGTTAAAAGTATTAGACAACATAGATGTGATTCGATCATTTTCA 369
Qy 1200 ATTACCGGATCGCTAAGTTCGATCGTGTAGATTTTGGATGGGAAAGCCTGCA 1259
Db 368 ATTGCTGATCGCCAAAACATGACTTATATCTGCTGCTGATTTGGATGGGAAAGCCGCA 309
Qy 1260 AAATTTGACATTACCTC-----TGTTGATTATGAGAAATGATTTATGTGATTCAG 1310
Db 308 AAGTTAGAATTTATTTCTATCGACAATGATGTTGTTGGAATTTCAATGCTCTTAGTAAA 249
Qy 1311 TCCAGGGATTTTGAAGAAGTGTGGAGATGGAGATGATCATTTGCCCTAAGATTCATATGAT 1370
Db 248 TCTAAGGATTTTGTATGGAGATTTAGAGATTTGCTGTTGTTGTTCTCTAAACTCGAATGAT 189
Qy 1371 GCATTTGCAAAATCTTTGAAGAAGGCTTTTTCCTTTGT 1410
Db 188 GCTTTGCTACTATATTCACCTCCCGGCTTAGCTTTCTAT 149

RESULT 12
LOCUS BE436185 542 bp mRNA EST 24-JUL-2000
DEFINITION EST407263 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG31A19, mRNA sequence.
ACCESSION BE436185
VERSION BE436185.1 GI:9434028
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 542)
AUTHORS Alcalá, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,
Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Ronning, C.M.,
Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley
, S.D.
Generation of ESTs from tomato fruit tissue, breaker stage
Unpublished (2000)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
LOCATION/Qualifiers
1. 542
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG31A19"

FEATURES
source
1. 542
/clone_lib="tomato breaker fruit, TIGR"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmCudapt; Site_1: EcoRI;
Site_2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."
```


FEATURES	source	Location/Qualifiers
	1. .659	
	/organism="Solanum tuberosum"	
	/cultivar="Kennebec"	
	/db_xref="taxon:4113"	
	/clone="CSTS20M23"	
	/clone_lib="cSTS"	
	/tissue_type="sprouting eyes from tubers"	
	/dev_stage="12-14 weeks post harvest"	
	/lab_host="SOLR"	
	/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."	
BASE COUNT	178 a 163 c 112 g 206 t	
ORIGIN		
Query Match	6.0%; Score 102.8; DB 155; Length 659;	
Best Local Similarity	53.4%; pred. No. 1.4e-16;	
Matches 310; Conservative	0; Mismatches 257; Indels 13; Gaps 4;	
Qy	34 TTGAAAATGCCAGTTACACACCATCTGACACACAGATGTGCAGTTATFCGCTACCGG	93
Db		
Qy	94 TAACATTTCTTCGATATCCCTGGTTGTCACITGATAGATGTCAGTCCTCTGTTTACG	153
Db		
Qy	154 ACTTTCCGTACCCAAAGAACACATTTCTTTGGACACTGTTATCCCTAACTTAAAGCCCTTT	213
Db		
Qy	209 AGCTACCAATTTCCAAACTCGATTTTCGTTCAACACCATTAATTCCTACTCTTAACATTCAC	268
Db		
Qy	214 TGTCTCTCACTCTAAAACACTACGTTCCGCTTAGCGGAAATTTGTTGATGCCGATCAAAAT	273
Db		
Qy	274 CGGCGGAATCCCGAAGTTTCAGTACTCCCCGTGATAGGGCGGACATCGATACTTTGATCG	333
Db		
Qy	334 TTGGCGAGCTGCACGAGATTTTGACTACCTTAAAGGTCATCAACTGGTAGATTTCCAATG	393
Db		
Qy	394 ATTTGCATGGCCCTTTTATGCTTATGCCACGGGTTATAAGGACCATGCAAGCATATAAG	453
Db		
Qy	454 TGATCCCGCTCGTAGCCGTGCAAGTAACCGTTTTCCTAACCGTGGCATACCGCT---GG	510
Db		
Qy	511 CTCTGACGGCACATCATTTCAATTCAGATGCTAAAAGTTTGTAAATGTTTCATCAATGCTT	570
Db		
Qy	571 GGGCTATATTAACAAAATTTGGGAAGACGCGGACTGTTT 610	
Db		
Qy	620 GGGGTTTACTCAACAAATTCGCGGTGATGAACAGTGCCTT 659	
Db		
RESULT 14		
AW930934		
LOCUS	AW930934 497 bp mRNA EST	30-MAY-2000
DEFINITION	EST356777 tomato fruit mature green, TAMU Lycopersicon esculentum	
VERSION	CNA clone cLEFA42M23 5', mRNA sequence.	
ACCESSION	AW930934	
KEYWORDS	AW930934.1 GI:8106335 EST.	

THIS PAGE BLANK (USPTO)

[illegible]

Db 2260 AATGTATACCTTCCAAATGCTGTTTGTGCTTTTACAAATAAATGATATGAACC 2319
QY 1676 TGTAAAAA 1703
Db 2320 TCAAAAAA 2347

RESULT 5
US-08-726-160-9
; Sequence 9, Application US/08726160
; Patent No. 5734016
; GENERAL INFORMATION:
; APPLICANT: LEVENS, DAVID L., DUNCAN,
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726.160
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/021.608
; FILING DATE: 22-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063US1
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2381
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
; ORIGINAL SOURCE:
; ORGANISM: Human
; CELL LINE: HL60
; FEATURE:
; OTHER INFORMATION: 470 bp variable
; OTHER INFORMATION: region where R is A or G.
US-08-726-160-9

Query Match 2.3%; Score 38.4; DB 1; Length 2381;
Best Local Similarity 64.8%; Pred. No. 0.22;
Matches 57; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY 1616 ATGTATAATGCCATTATATCTTCCAAATGCTGTTTGTGCTTTTACAAATAAATGATATGAACC 2319
Db 2260 AATGTATACCTTCCAAATGCTGTTTGTGCTTTTACAAATAAATGATATGAACC 2319
QY 1676 TGTAAAAA 1703

Db 2320 TCAAAAAA 2347
RESULT 6
PCT-US94-01782-9
; Sequence 9, Application PC/TUS9401782
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES
; APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF
; APPLICANT: HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01782
; FILING DATE: 22-FEB-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/021.608
; FILING DATE: 22-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063PCT
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2381
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
; ORIGINAL SOURCE:
; ORGANISM: Human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE: HL60
; ORGANELLE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: 470 bp variable
; OTHER INFORMATION: region where R is A or G.
PCT-US94-01782-9

Query Match 2.3%; Score 38.4; DB 5; Length 2381;
Best Local Similarity 64.8%; Pred. No. 0.22;


```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..183
; US-09-065-474-102

Query Match      2.1%; Score 36.6; DB 3; Length 184;
Best Local Similarity 57.4%; Pred. No. 0.21;
Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1042 CACATAAAGAGTTAGTTGGGGATAAAGGGCTTCTGTTGCAGTTGCCAGCTATTGGAGAA 1101
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26 CACAAGAACCCGTTTTTTTCAGCAAGTAGGGTATAATGGGACACCTGGAGATCCAGAAGAAC 85
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1102 CCATGTAAAGAGGTTGCACAACGAAAAAGCGCTTCTTGCAGAGTGCAAAAACTTG 1156
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 86 CCAAGAAGACAGTCATTCCACGAAGAGGAGGAGTTGCAGTTTCAGAAATTCG 140
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-09-065-474-103/C
; Sequence 103, Application US/09065474
; Patent No. 6063599
; GENERAL INFORMATION:
; APPLICANT: Tang, Liang

```

APPLICANT: BERNI, E. SCOT
TITLE OF INVENTION: DIROFILARIA AND BRUGIA ANKYRIN
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES, AND
TITLE OF INVENTION: USES THEREOF

```

CORRESPONDENCE ADDRESS:
ADDRESS: Carol Talkington Verser, Ph.D.
ADDRESSEE: Heska Corporation
STREET: 1825 Sharp Point Drive
CITY: Fort Collins
STATE: Colorado
COUNTRY: USA
ZIP: 80525
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Wordperfect for Windows, Version 7.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/065,474
FILING DATE: 24-APR-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Verser, Carol Talkington
REGISTRATION NUMBER: 37,459
REFERENCE/DOCKET NUMBER: HW-5-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 970/493-7272
TELEFAX: 970/484-9505
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 184 nucleotides
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-09-065-474-103

Query Match 2.18; Score 36.6; DB 3; Length 184;
Best Local Similarity 57.4%; Pred. No. 0.21;
Matches 66; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 1042 CACATAAAGAGTTAGTTGGGGATAAAGGGCTTCTTTGTCAGTTGCAGCTATTTCGAGAAG 1101
|||||

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 13:54:25 ; Search time 445.49 Seconds
(without alignments)
2400.315 Million cell updates/sec

Title: US-08-894-356C-1

Perfect score: 1703

Sequence: 1 TCATTATGAGCAATCCAA.....AAAAAAAAAAAAAAAAAAAA 1703

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

1: /SIDSL/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1698.2	99.7	1703	17 AAT37308 Aromatic acyl tran
2	202.4	11.9	1508	17 AAT37312 Aromatic acyl tran
3	193	11.3	1518	17 AAT37313 Aromatic acyl tran
4	186	10.9	1622	17 AAT37309 Aromatic acyl tran
5	172.2	10.1	1479	17 AAT37311 Aromatic acyl tran
6	110	6.5	936	22 AAF58252 Oligonucleotide D1
7	110	6.5	936	22 AAF58254 Oligonucleotide D1
8	110	6.5	936	22 AAF58257 Oligonucleotide D1
9	110	6.5	936	22 AAF58259 Oligonucleotide D2
10	110	6.5	936	22 AAF58262 Oligonucleotide D2
11	110	6.5	938	22 AAF58255 Oligonucleotide D1

12	106.8	6.3	936	22 AAF58252 Oligonucleotide D1
13	106.8	6.3	936	22 AAF58254 Oligonucleotide D1
14	106.8	6.3	936	22 AAF58257 Oligonucleotide D1
15	106.8	6.3	936	22 AAF58259 Oligonucleotide D2
16	106.8	6.3	936	22 AAF58262 Oligonucleotide D2
17	106.8	6.3	938	22 AAF58255 Oligonucleotide D1
18	81	4.8	1648	21 AAC64785 Lemon acyl transe
19	53.6	3.1	1379	21 AAC64852 Arabidopsis thalia
20	45.4	2.7	1502	21 AAC49421 Arabidopsis thalia
21	45.4	2.7	1504	21 AAC33536 Arabidopsis thalia
22	45	2.6	244	22 AAF58238 Oligonucleotide D1
23	44.2	2.6	244	22 AAF58238 Oligonucleotide D1
24	43	2.5	1240	21 AAC53227 Arabidopsis thalia
25	42.4	2.5	1486	21 AAC98137 Human colon cancer
26	41.8	2.5	2222	21 AAA27236 Human MEK1 DNA. H
27	40	2.3	408	22 AAF17559 Human breast cance
28	40	2.3	2691	21 AAC79685 Pinus radiata cell
29	39.8	2.3	959	22 AAC66940 Proteasome subunit
30	39.8	2.3	3581	15 AAQ44391 Sequence of murine
31	39	2.3	340	20 AAX51737 DNA encoding a hum
32	39	2.3	1373	19 AAV52969 Human G1 protein 1
33	39	2.3	1701	21 AAZ33336 Human secreted pro
34	39	2.3	5713	21 AAC76360 Human ORFX ORF1915
35	39	2.3	7122	22 AAF31100 Methionine synthas
36	39	2.3	7122	22 AAC91207 Human methionine s
37	38.8	2.3	1544	21 AAA46483 cDNA encoding a fi
38	38.8	2.3	1670	22 AAF24992 Nucleotide sequenc
39	38.6	2.3	932	21 AAZ51920 N. tabacum TOB-F12
40	38.6	2.3	1106	21 AAA29167 Rice 4-hydroxyphen
41	38.6	2.3	4292	21 AAF21792 Human breast and o
42	38.4	2.3	2381	15 AAO68910 Human c-myc far up
43	38.4	2.3	2384	15 AAO68909 Human c-myc far up
44	38.2	2.2	1570	21 AAC76262 Human ORFX ORF1817
45	38.2	2.2	1820	20 AAZ00434 Human secreted pro

ALIGNMENTS

RESULT 1

AAT37308

ID AAT37308 standard; cDNA to mRNA; 1703 BP.

AC AAT37308;

DT 06-FEB-1997 (first entry)

XX Aromatic acyl transferase coding sequence.

DE Aromatic acyl transferase: transformation; anthocyanin pigment;

XX plants; acylation: colour; tone; colouration; colour change;

KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;

KW Scenecio cruentus; Lavandula angustifolia; ds.

OS Gentiana triflora var. japonica (Clone pcAT4).

FH Key Location/Qualifiers

FT CDS 6..1415

FT /*tag= a

FT /product= Aromatic acyl transferase.

XX WO9625500-A1.

XX 22-AUG-1996.

XX 16-FEB-1996; 96WO-JP00348.

XX 30-JAN-1996; 96JP-0046534.

XX 17-FEB-1995; 95JP-0067159.

XX 29-JUN-1995; 95JP-0196915.

XX (SUNR) SUNTORY LTD.

PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
PI Nakao M, Tanaka Y, Yonekura K;
XX WPI: 1996-393401/39.
DR P-PSDB; AAW04722.
XX DNA coding for aromatic acyl transferase - for transforming plants
PT which produce anthocyanin pigments and thus altering colour tone,
PT e.g. of flowers
XX
PS Claim 4; Page 53-57; 94pp; Japanese.
XX Vectors containing DNA fragments encoding proteins of plant origin
CC with aromatic acyl transferase activity may be used to transform
CC plants which produce anthocyanin pigments. The aromatic acyl
CC transferase acylates the pigments in the flower resulting in colour
CC tone changes and allowing new colourations to be produced. Six
CC specific DNA sequences encoding aromatic acyl transferase from
CC different plants are described in AAT37308-T37313.
XX
SQ Sequence 1703 BP; 512 A; 353 C; 356 G; 482 T; 0 other;

Query Match 99.7%; Score 1698.2; DB 17; Length 1703;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1700; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCATTATGGAGCAATCCAAATGTTGAAGTCTTGTGAAAATGCCAAGTTACACCACCAT 60
DB 1 tcattatggagcaaatccaaatggtgaagttcttgaataatgccaaatgacacacacac 60
QY 61 CTGACAAACAGAGTGCAGTATTCGTACCGGTAAACATCTTCGATATCCCGTGGTTGC 120
DB 61 ctgacaaacagagtgcagtattcgtacccggttaacattcttcgatccctcgtgtgc 120
QY 121 ACTTGAATAAGATCAGTCCCTCTGTTTACGACTTTCGCTACCCAGAACACATTCT 180
DB 121 acttgaataagatcagtcctctctgtttacgactttccgacttcgacccaagaacattct 180
QY 181 TGGACACTGTTATCCCAATCTTAAGCCCTCTTGTCTCTACATCTAAACACTACGTT 240
DB 181 tggacactgttatcccaatcttaagccctcttgtctctcactctaaacactacacct 240
QY 241 CGCTTAGCGGAAATTTGTTGATCGCGATCAATCGCGGAAATCCCGAAGTTTCAGTACT 300
DB 241 cgccttagcggaaatttggttgatacggatcaaatcgcggaatcgcggaatcttcagttact 300
QY 301 CCCGTGATAGGGCGACTCGATACCTTGAATTCATGATTTGCATGGCTTTTATGTTATGC 420
DB 301 cccgtgatagggcgactcgataccttgaattcattgcatggtcttcttattgtatgc 420
QY 421 CACGGGTATTAAGGACCATGCAACACTATAAGTGATCCCGCTCGTAGCCGTGCAAGTAA 480
DB 421 cacgggttataaggacccatgcaaacactataaagtgatcccgctcgtagccgtgcaagtaa 480
QY 481 CCGTTTTTCTTAACCGGTGATACCGCTGGCTCTGAGGGCACATCATTCATTAATTCAGATG 540
DB 481 ccgttttcttaacccggtgatatccgctggctctgagcgacacatcatcattcaatgcatg 540
QY 541 CTAAGAGTTTGTAAATGTTATCAATGCTTTGGGCTATATTAACAAATTTGGGAAAGACG 600
DB 541 ctaagagtttgttaattgtatcaatgctttgggctattataacaaatttgggaaagacg 600
QY 601 CGGACTTTGTTCCGGAATCTTCTCCATCTTTCGATAGATCATTAATCAAGATCTGT 660
DB 601 cggactttgttccggaaattcttctccattcttcgatagatcagataatcaagatctgt 660
QY 661 ATGCCCTAGAGGAACATTTTGGAAACGAAATGCAAGATGTTCTTGAATGTTCTCTAGAT 720
DB 661 atgcccttagaggaacatTTTGGAAACGAAATGCAAGATGTTCTTGAATGTTCTCTAGAT

Best Local Similarity	1.1%;	Fied. NO: 8.7E-20;
Matches	8;	Conservative 434; Mismatches 264; Indels 0; Gaps 0;

DD 222 WWWWWW

[illegible]

Db 517 www..... 576
Qy 1538 ATAAATCCAGAGTCGAATATACACCGTTCTCTCGTAAAGTTGAACCTCACACCTGA 1597
Db 577 www..... 636
Qy 1598 CATGGTGTACGATAGGTATTGTATAATGCCATTATATCTTCCATTAAGTATCCTATGC 1657
Db 637 www..... 696
Qy 1658 AATAGAGAACATGTATGTGTTAAAAA..... 1703
Db 697 www..... 742

RESULT 13
AAF58254
ID AAF58254 standard; DNA; 936 BP.
XX AC AAF58254;
XX DT 24-APR-2001 (first entry)
XX DE Oligonucleotide D1875.
XX KW Electron-transfer group; ETM; mismatch; genotyping;
XX KW gene expression; ss.
XX OS Synthetic.
XX PN WO200107665-A2.
XX PD 01-FEB-2001.
XX PF 26-JUL-2000; 2000WO-US20476.
XX PR 26-JUL-1999; 99US-0145695.
XX PR 17-MAR-2000; 2000US-0190259.
XX PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX PI Umek RM;
XX DR WPI; 2001-159728/16.
XX PT Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface -

Example 6; Page 127; 159pp; English.
The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.
Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 6.3%; Score 106.8; DB 22; Length 936;
Best Local Similarity 0.8%; Pred. No. 6.3e-19;
Matches 6; Conservative 434; Mismatches 266; Indels 0; Gaps 0;

Qy 998 TAACACTTTGGCAACTGTCTTGGCTCATCGTTGCAAAAGAACACATAAAGATTAGT 1057
Db 37 www..... 96
Qy 1058 TGGGGATAAAGGGCTTCTGTTCGAGTTGCAGCTATTGGAGAGCCATTGAAAGAGGTT 1117
Db 97 www..... 156

Qy 1118 GCACAAGAAAAAGCGCTTCTTCGAGATCAAAAACTTGGTTATCGGAATCAATGAAT 1177
Db 157 www..... 216
Qy 1178 CCCTTCAAAAGATTCTCGGGATTACCGGATTCGCCCTAAGTTCGATTCGATGGTGTA 1237
Db 217 www..... 276
Qy 1238 TTTTGGATGGGAAGCCTGCAAAATTTGACATTACCTCTGTTGATTATGCAGATTGAT 1297
Db 277 www..... 336
Qy 1298 TTATGTGATTTCAGTCCAGGATTTTGAAGAGGTGTGAGATTGGAGTATCATTCCTAA 1357
Db 337 www..... 396
Qy 1358 GATTCATATGATGATTCGAAAAATCTTTGAAGAGGCTTTTGTCTTTGTCTATAGTC 1417
Db 397 www..... 456
Qy 1418 TCTTTAATAGAACCATATTGCTGCAATAAAGTACCAAGTCTCTTACTAACACTACACCA 1477
Db 457 www..... 516
Qy 1478 AACCTACTTTTCGAGCGGGAACACACAGAGGTTTCACTACTAGAGGTTGACTTC 1537
Db 517 www..... 576
Qy 1538 ATAAATCCAGAGTCAATATACACCGTTGCTCTGTAAGAGTTGAACCTCACACCTGA 1597
Db 577 www..... 636
Qy 1598 CATGGTGTACGATAGGTATTGTATAATGCCATTATATCTTCCATTAAGTATCCTATGC 1657
Db 637 www..... 696
Qy 1658 AATAGAGAACATGTATGTGTTAAAAA..... 1703
Db 697 www..... 742

RESULT 14
AAF58257
ID AAF58257 standard; DNA; 936 BP.
XX AC AAF58257;
XX DT 24-APR-2001 (first entry)
XX DE Oligonucleotide D1954.
XX KW Electron-transfer group; ETM; mismatch; genotyping;
XX KW gene expression; ss.
XX OS Synthetic.
XX PN WO200107665-A2.
XX PD 01-FEB-2001.
XX PF 26-JUL-2000; 2000WO-US20476.
XX PR 26-JUL-1999; 99US-0145695.
XX PR 17-MAR-2000; 2000US-0190259.
XX PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX PI Umek RM;
XX DR WPI; 2001-159728/16.
XX PT Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:04:15 ; Search time 168.74 seconds
(without alignments)
1800.666 Million cell updates/sec

Title: US-08-894-356C-3
Perfect score: 1605
Sequence: 1 TGTCGACGAATCCATTCA.....TTAAAAA.....1605

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/lna/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	67	4.2	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	45	2.8	2852	3 US-09-027-137-2	Sequence 2, Appl1
C 3	44.8	2.8	1474	4 US-08-821-994-64	Sequence 64, Appl
C 4	43.8	2.7	3157	6 5198347-3	Patent No. 5198347
C 5	43.6	2.7	1582	3 US-08-545-196B-10	Sequence 10, Appl
C 6	43.6	2.7	1582	3 US-08-545-196B-12	Sequence 12, Appl
C 7	43.6	2.7	1882	4 US-09-370-253-1	Sequence 1, Appl1
C 8	43.4	2.7	1046	1 US-08-361-467B-4	Sequence 4, Appl1
C 9	43.4	2.7	1332	2 US-08-484-332C-4	Sequence 4, Appl1
C 10	43.2	2.7	1332	2 US-09-057-762-1	Sequence 1, Appl1
C 11	43.2	2.7	1332	3 US-08-326-119A-1	Sequence 1, Appl1
C 12	41.8	2.6	3337	1 US-08-072-610-1	Sequence 1, Appl1
C 13	41.8	2.6	3337	2 US-08-719-822B-1	Sequence 1, Appl1
C 14	41.8	2.6	3337	4 US-09-092-458-1	Sequence 1, Appl1
C 15	41.6	2.6	581	2 US-08-557-309B-22	Sequence 22, Appl
C 16	41.6	2.6	581	3 US-08-834-306-22	Sequence 22, Appl
C 17	41.6	2.6	581	4 US-08-993-674A-22	Sequence 22, Appl
C 18	41.6	2.6	5852	1 US-07-867-106-2	Sequence 2, Appl1
C 19	40.8	2.5	746	4 US-09-013-810-1	Sequence 1, Appl1
C 20	40.6	2.5	240	1 US-08-628-417-6	Sequence 6, Appl1
C 21	40.4	2.5	3527	2 US-08-909-965C-7	Sequence 7, Appl1
C 22	40	2.5	140	1 US-08-628-417-5	Sequence 5, Appl1
C 23	40	2.5	1798	2 US-08-557-128-12	Sequence 12, Appl
C 24	40	2.5	2295	1 US-08-375-300-3	Sequence 3, Appl1
C 25	40	2.5	2295	3 US-09-177-431-3	Sequence 3, Appl1
C 26	40	2.5	2295	5 PCT-US95-16930-3	Sequence 3, Appl1
C 27	40	2.5	4080	1 US-08-375-300-1	Sequence 1, Appl1

28	2.5	4080	3	US-09-177-431-1	Sequence 1, Appl1
29	40	2.5	4080	5 PCT-US95-16930-1	Sequence 1, Appl1
30	39.8	2.5	358	2 US-08-721-488-9	Sequence 9, Appl1
31	39.4	2.5	923	4 US-09-004-731-66	Sequence 66, Appl1
32	39.4	2.5	923	4 US-08-749-699-66	Sequence 66, Appl1
C 33	39	2.4	1736	3 US-09-182-816-22	Sequence 22, Appl1
C 34	39	2.4	1736	3 US-09-182-816-24	Sequence 24, Appl1
C 35	39	2.4	1736	3 US-09-471-528-22	Sequence 22, Appl1
C 36	39	2.4	1736	3 US-09-471-528-24	Sequence 24, Appl1
C 37	38.8	2.4	8920	2 US-08-446-855A-1	Sequence 1, Appl1
C 38	38.8	2.4	8920	4 US-09-150-741-1	Sequence 1, Appl1
39	38.6	2.4	6877	1 US-08-347-340-1	Sequence 1, Appl1
40	38.4	2.4	260	2 US-08-520-678A-29	Sequence 29, Appl1
C 41	38.4	2.4	2230	4 US-08-378-313-24	Sequence 24, Appl1
42	38.2	2.4	923	3 US-08-906-769-144	Sequence 144, App
43	38.2	2.4	923	3 US-08-906-616-144	Sequence 144, App
44	38.2	2.4	923	3 US-08-639-075A-144	Sequence 144, App
45	38.2	2.4	923	4 US-09-012-431-144	Sequence 144, App

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFELINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22131-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DCS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; TOPOLOGY: single
; STRANDEDNESS: linear
; IMMEDIATE SOURCE:
; CLONE: pTZapt-Fls
US-08-232-463-14

```
Query Match          4.2%; Score 67; DB 1; Length 7218;
Best Local Similarity 2.0%; Pred. No. 2.7e-08;
Matches 7; Conservative 221; Mismatches 121; Indels 0; Gaps 0;

Qy 1044 TTGTTACTACCAAGCCGATCGAGTTCGCTGGTGGGATGATACACAAGCGATCGTGA 1103
      |||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1442 TTGTTACTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1383
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 1104 GCATGACCTAAGGCCATTGATGAAGAAACAAGAGTGGGAGAGCAACCCGAGATCTT 1163
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1382 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1323
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 1164 TCAGTACAAAGATCGTGAGTCAACTGTGCTGTGGAAGTTCGCCAAGGTTCAAGGT 1223
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1322 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1263
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 1224 TTACGACGTGATTGATGGGNAACCCAGAGCTGTGAGGAGTGTTCGAACAATAG 1283
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1262 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1203
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 1284 GTTGTGATGAATGGTGTATTGTACCAAGCAAAATGGAGGAAGCATTTGATGTGA 1343
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1202 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1143
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 1344 GATTAGTTGGAAGCAAAATGCTATGGAGAGGTTCGAGAAAGATAAAGAG 1392
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1142 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1094
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
US-09-027-137-2/c
; Sequence 2, Application US/09027137
; Patent No. 6013450
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: CAF1-RELATED PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/027,137
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0476 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2852 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT16
```

```
; CLONE: 2229466
US-09-027-137-2

Query Match          2.8%; Score 45; DB 3; Length 2852;
Best Local Similarity 56.4%; Pred. No. 0.014;
Matches 84; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 1453 GAGCTGCTATGACACTCACATGATGATGTTTATTTTGGAGGGTCTCTTCCTTTTA 1512
      |||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2599 GAGCGGCCCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2540
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 1513 TTGTTTCTATGTTTCTTTCTTTGATGATGATGAGAGAAACCGAGTATAAAGAAATA 1572
      |||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2539 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 2480
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 1573 ATGTTTTCAGTTTATTAATAAAAAAAAAAAAAA 1601
      | ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2479 TTTATTTTACAAAAATTAATAACATAAATA 2451

RESULT 3
US-08-821-994-64/c
; Sequence 64, Application US/08821994A
; Patent No. 6228643
; GENERAL INFORMATION:
; APPLICANT: Greenland, Andrew J
; APPLICANT: Thomas, Didier Rp
; APPLICANT: Jepson, Ian
; TITLE OF INVENTION: Promoters
; FILE REFERENCE: PPD 50108
; CURRENT APPLICATION NUMBER: US/08/821,994A
; CURRENT FILING DATE: 1997-03-22
; EARLIER APPLICATION NUMBER: PCT/GB97/00729
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: GB 9606062.9
; EARLIER FILING DATE: 1996-03-22
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Brassica napus
US-08-821-994-64

Query Match          2.8%; Score 44.8; DB 4; Length 1474;
Best Local Similarity 59.4%; Pred. No. 0.011;
Matches 76; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 1475 TGTATGTTTATTTTGGAGGGTCTCTTCTTTTATTTGTTTCTATGTTTCTTTTCTTT 1534
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1448 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 1389
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 1535 CTTGTACGTTATGAGAGAAACCGAGTATAAAGGAATATGTTTTCAGTTATTAATAAAA 1594
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1388 TTTTCTTTTCTTTTAAAGAGAAATATATCTTTTATATATCTCTCTTAATACATAAACAC 1329
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 1595 AAAAAAA 1602
      ||| ||
Db 1328 CAAATCAA 1321

RESULT 4
5198347-3
; Patent No. 5198347
; APPLICANT: Miller, Louis H.; ADAMS, JOHN H.; KASLOW,
; DAVIC C.; FANG, XIANGDOUG
; TITLE OF INVENTION: DNA ENCODING PLASMODIUM VIVAX AND
; PLASMODIUM KNOWLESII DUPEY RECEPTOR
; NUMBER OF SEQUENCES: 27
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/554,837
```


; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370,253
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1882
; TYPE: DNA
; ORGANISM: Hordeum vulgare
US-09-370-253-1

Query Match 2.7%; Score 43.6; DB 4; Length 1882;
Best Local Similarity 57.2%; Pred. No. 0.027;
Matches 79; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
Qy 1444 TTTATTTATGAGCTGCTATGACTCACATGCATGTTATTTTGGAGGGTTCT 1503
Db 1879 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1820
Qy 1504 TTCCTTTATGTTCTTCATGTTTCTTCTGCTAGCTTATGAAGAAACCGAGTAT 1563
Db 1819 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1760
Qy 1564 AAGAGATAATGTTTCA 1581
Db 1759 TCGGAATAATCTATGCA 1742

RESULT 8
US-08-361-467B-4/c
; Sequence 4, Application US/08361467B
; Patent No. 5633441
; GENERAL INFORMATION:
; APPLICANT: De Greef, Willy
; APPLICANT: Van Emmelo, John
; APPLICANT: De Oliveria, Dulce E.
; APPLICANT: De Souza, Maria-Helena
; APPLICANT: Van Montagu, Marc
; TITLE OF INVENTION: PLANTS WITH MODIFIED FLOWERS, SEEDS OR
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/361,467B
; FILING DATE: 22-DEC-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,492
; FILING DATE: 04-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP90/01275
; FILING DATE: 01-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 89 402 224.3
; FILING DATE: 04-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Schulman, Robert M.

; REGISTRATION NUMBER: 31,196
; REFERENCE/DOCKET NUMBER: 010830-027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1046 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; IMMEDIATE SOURCE:
; CLONE: 3C9
US-08-361-467B-4

Query Match 2.7%; Score 43.4; DB 1; Length 1046;
Best Local Similarity 57.9%; Pred. No. 0.023;
Matches 77; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
Qy 1473 CATGTAATGTTTATTTTTCGAGGGTTCCTTCCTTTATGTTTCTATGTTTCT 1532
Db 1042 CCGCATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 983
Qy 1533 TTCTGTAGCTTATGAAGAAACCGAGTATAAAGGAATAATGTTTCAGTTATTAAAA 1592
Db 982 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 923
Qy 1593 AAAAAAATAAAAA 1605
Db 922 GAGAGAAACAAAA 910

RESULT 9
US-08-484-332C-4/c
; Sequence 4, Application US/08484332C
; Patent No. 5767374
; GENERAL INFORMATION:
; APPLICANT: De Greef, Willy
; APPLICANT: Van Emmelo, John
; APPLICANT: De Oliveria, Dulce E.
; APPLICANT: De Souza, Maria-Helena
; APPLICANT: Van Montagu, Marc
; TITLE OF INVENTION: PLANTS WITH MODIFIED FLOWERS, SEEDS OR
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,332C
; FILING DATE: 7-JUNE-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/361,467
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681,492
; FILING DATE: 04-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP90/01275
; FILING DATE: 01-AUG-1990
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: EP 89 402 224.3
FILING DATE: 04-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Schulman, Robert M.
REGISTRATION NUMBER: 31,196
REFERENCE/DOCKET NUMBER: 010830-093
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1046 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
IMMEDIATE SOURCE:
CLONE: 3C9
US-08-484-332C-4

Query Match 2.7%; Score 43.4; DB 1; Length 1046;
Best Local Similarity 57.9%; Pred. No. 0.023;
Matches 77; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 1473 CATGTATGTTTATTTTGGAGGGTCTTTCCTTTTATTTCTTCTATGTTTCT 1532

Db 1042 CTGCGATTT 983

QY 1533 TTCTTCTAGCTTAAGAAGAACCGAGTATAAAGGAATAAGTTTTCAGTTATTAAAA 1592

Db 982 TTTTTTTTTTTTTTTTGAAGAAAACTTATAAATCCATTTACTTGTGGTGTGATA 923

QY 1593 AAAAAAAAAAAAAA 1605
Db 922 GAGAGAAACAAAA 910

RESULT 10

US-09-057-762-1/c
Sequence 1, Application US/09057762
Patent No. 5879909

GENERAL INFORMATION:
APPLICANT: PERL, ANDRAS
TITLE OF INVENTION: HUMAN TRANSALDOLASE: AN AUTOANTIGEN WITH
TITLE OF INVENTION: A FUNCTION IN METABOLISM
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812

COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/057,762
FILING DATE:
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/326,119
FILING DATE: 19-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: LIVNAT, SHMUEL
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: 280932000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168

TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1332 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 57..1064
US-09-057-762-1

Query Match 2.7%; Score 43.2; DB 2; Length 1332;
Best Local Similarity 57.4%; Pred. No. 0.029;
Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1470 ATGCATGATGTTTATTTTGGAGGGTCTTTCCTTTTATTTGTTTCTATGTTTT 1529
Db 1330 ATTCCCTTT 1271

QY 1530 TCTTTCTGTACGTTATGAAGAACCGAGTATAAAGGAATAAGTTTTCAGTTATTAA 1589
Db 1270 TTTTTTTTTTTTTTTTTTTTGAACACAGCAGGAAAGTGCTGCTTAATGTTATTAG 1211

QY 1590 AAAAAAAAAAAAAA 1605
Db 1210 GCAAAATTTTACATAA 1195

RESULT 11

US-08-326-119A-1/c
Sequence 1, Application US/08326119A
Patent No. 6018021

GENERAL INFORMATION:
APPLICANT: PERL, ANDRAS
TITLE OF INVENTION: HUMAN TRANSALDOLASE: AN AUTOANTIGEN WITH
TITLE OF INVENTION: A FUNCTION IN METABOLISM
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20006-1812

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/326,119A
FILING DATE: 19-OCT-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: LIVNAT, SHMUEL
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: 280932000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 822-0168

TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1332 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 57..1064
US-08-326-119A-1

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:00:43 ; Search time 5816.79 Seconds
(without alignments)
2608.282 Million cell updates/sec

Title: US-08-894-356C-3
Perfect score: 1605
Sequence: 1 TGTGCGAGGAATCATTCA.....TTAAAAAAAAAAAAAAAAAAAA 1605

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estrol:*
83: em_estrol2:*
84: em_estrol3:*
85: em_estrol4:*
86: em_estrol5:*
87: em_estrol6:*
88: em_estrol7:*
89: em_estrol8:*
90: em_estrol9:*
91: em_estrol10:*
92: em_estrol11:*
93: em_estrol12:*
94: em_estrol13:*
95: em_estrol14:*
96: em_estrol15:*
97: em_estrol16:*
98: em_estrol17:*
99: em_estrol18:*
100: em_estrol19:*
101: em_estrol20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	428	26.7	632	141	BE919635	BE919635 EST423404
2	420.2	26.2	657	119	AW649974	AW649974 EST328428
3	406.4	25.3	649	174	BG126859	BG126859 EST472505
4	404.2	25.2	642	143	BF051093	BF051093 EST436268
5	403	25.1	588	141	BE919634	BE919634 EST423403
6	401.8	25.0	614	122	AW928744	AW928744 EST337532
7	386.8	24.1	651	174	BE919635	BE919635 EST472505
8	371.4	23.1	582	118	AW618433	AW618433 EST320419
9	363.6	22.7	569	142	BE920362	BE920362 EST424131
10	355.4	22.1	564	119	AW039511	AW039511 EST281792
11	355.2	22.1	544	119	AW649611	AW649611 EST328065
12	353.8	22.0	556	143	BF053244	BF053244 EST438474
13	345	21.5	786	146	BF275133	BF275133 GA_EB002
14	343.8	21.4	563	110	AW041172	AW041172 EST284036
15	332.2	20.7	545	118	AW617090	AW617090 EST323501
16	331.2	20.6	531	167	BE458676	BE458676 EST413968
17	326	20.3	517	173	BG096122	BG096122 EST450641
18	322.8	20.1	508	142	BE924264	BE924264 EST428033
19	321	20.0	535	166	BE353229	BE353229 EST400361
20	315.2	19.6	870	153	BG446054	BG446054 GA_Ea003
21	315.2	19.6	878	77	AW731285	AW731285 GA_Ea003
22	314.4	19.6	495	143	BF054047	BF054047 EST439277
23	306.6	19.1	573	115	AW939481	AW939481 EST309981
24	306.2	19.1	695	76	AW668525	AW668525 GA_Ea001
25	305.6	19.0	510	118	AW617259	AW617259 EST323670
26	305.4	19.0	687	153	BG441719	BG441719 GA_Ea001
27	301.6	18.8	882	76	AW728139	AW728139 GA_Ea001
28	301.2	18.8	624	103	AW189979	AW189979 EST2404.Y
29	300.4	18.7	522	120	AW735992	AW735992 EST336760
30	300.2	18.7	874	153	BG441701	BG441701 GA_Ea001
31	298.2	18.6	594	24	AW130615	AW130615 BNGLH1741
32	296.6	18.5	501	118	AW617396	AW617396 EST323807
33	296.4	18.5	477	143	BF053972	BF053972 EST439202
34	294.4	18.3	629	24	AW1727325	AW1727325 BNGLH1775
35	294.2	18.3	508	118	AW616992	AW616992 EST323403
36	288.6	18.0	537	24	AW130339	AW130339 BNGLH1668
37	287.2	17.9	652	119	AW690825	AW690825 NF035D15
38	283.4	17.7	454	115	AW399665	AW399665 EST310165
39	283.4	17.7	511	103	AW1901068	AW1901068 sc19b1.Y
40	281.6	17.5	640	24	AW130449	AW130449 BNGLH1720
41	279.8	17.4	660	146	BF268244	BF268244 GA_EB000
42	274.6	17.1	641	24	AW130451	AW130451 BNGLH1721
43	273.8	17.1	658	146	BF278351	BF278351 GA_EB003
44	271.6	16.9	474	122	AW934606	AW934606 EST353498
45	268.4	16.7	611	24	AW130368	AW130368 BNGLH1675

ALIGNMENTS

RESULT	1	BE919635	632 bp	mRNA	EST	02-OCT-2000
LOCUS	BE919635/c	EST423404	potato leaves and petioles	Solanum tuberosum	cDNA clone	
DEFINITION		CSTB2C8 3' sequence, mRNA sequence.				
ACCESSION		BE919635				
VERSION		BE919635.1	GI:10445711			
KEYWORDS		EST.				
SOURCE		potato.				
ORGANISM		Solanum tuberosum				
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
		Asteridae; euasterids I; Solanales; Solanaceae; Solanum.				
REFERENCE		1 (bases 1 to 632)				
AUTHORS		van der Hoeven, R.S., Bezzerides, J., Holt, I.E., Liang, F., Cho, J.,				
		Utterback, T., Hansen, C.L., Doan, B., Bougri, O., Buell, C.R., Ronning				
		, C.M., Fry, W.E., Tanksley, S.D. and Baker, B.				

TITLE
JOURNAL
COMMENT

Generation of ESTs from potato leaves and petioles
Unpublished (2000):
Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
Location/Qualifiers

FEATURES
source

1. 632
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB2C8"
/tissue_type="potato leaves and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen." 183 a 181 c 95 g 173 t

BASE COUNT
ORIGIN

Query Match 26.7%; Score 428; DB 141; Length 632;
Best Local Similarity 81.2%; Pred. No. 3.6e-95;
Matches 497; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY	843	CTCAGGAGAGCGCCATTCTCCACATTTCCAATCTCTCCGACACAGCTGGCTAGCCGT	902
DB	631	CTCAATAGCCGCTGCTCTTCTTCTTCCAGTCCCTCTCCGACACAGCTGGCTAGCTGT	572
QY	903	CACAGCTGGCGGCCAACTCAACCGGAGACTACACTGTGTACACTGTTTGTGCTGATTG	962
DB	571	CACAGCTGGCGGCCAACTCAACCGGAGACTACACTGTGTACACTGTTTGTGCTGATTG	512
QY	963	CAGGAAAGGGTTGATCTCTCAATGCCAGAAAGTTACTTCGCAACCTAATTCAGCGCAT	1022
DB	511	CCGGAAGAGTGGATCTCTCAATGCCAGAAAGTTACTTCGCAACCTAATTCAGCGCAT	452
QY	1023	TTTTCAGTGCAGCGCGGAGTCTTGTACTAGCAAGCCGATCGAGTTCGCTGGTGGAT	1082
DB	451	TTTTCAGTGCAGCGCGGAGTCTTGTACTAGCAAGCCGATCGAGTTCGCTGGTGGAT	392
QY	1083	GATACAAAGCGATCGTGAACGATGACCTAAGGCCATTGATGAAGAAACAGAGATG	1142
DB	391	GATTGAGCAAGCAATGTGAATCAGCATCGAAGCGGATTTGATGAAGAAACAGAGATG	332
QY	1143	GGAGAGCAACCGGAGATCTTTTCACTACAAAGATGCTGGAGTGAATTTGCTGTTGG	1202
DB	331	GGAAAGTAGTCCGAGATATTTGAGTACAAAGATGCGGAGTGAATTTGCTGCTGGG	272
QY	1203	AAGTTCCGCAAGTTCAAGTTTACGACGTTTGGATTTGGATGGGAAACCCAGAGATGT	1262
DB	271	AAGTTCCGCAAGTTTCAAGTTTACGACGTTTGGATTTGGATGGGAAACCCAGAGATGT	212
QY	1263	GAGGAGTGGTTCGACAAATAGCTTTGATGGAATGTTGATGATGATGATGATGATGATG	1322
DB	211	GAGAAGTGGATTTGAACAATAGCTTTGATGGAATGTTGATGATGATGATGATGATGATG	152
QY	1323	AGGAAGCAATTTGATGGAATAGCTTTGGAACCAATGCTATGAGAGGTTGGAGAA	1382
DB	151	AGGAAGCAATTTGATGGAATAGCTTTGGAACCAATGCTATGAGAGGTTGGAGAA	92
QY	1383	AGATAAGAGTTCCTCATGGAACCTGCTTAATTTGCTTGGACTCAACTGGCTACA	1442
DB	91	GGATAAGAGTTCCTCATGGAACCTGCTTAATTTGCTTGGACTCAACTGGCTACA	32
QY	1443	CTTTATTTATGA	1454
DB	31	TTGTTTTTTTAA	20

```

RESULT 2
AW649974
LOCUS AW649974 657 bp mRNA EST 04-APR-2000
DEFINITION EST328428 tomato germinating seedlings, TAMU Lycopersicon
esculentum cDNA clone cLE11l6 5', mRNA sequence.
ACCESSION AW649974
VERSION AW649974.1 GI:7411212
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 657)
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S, Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Giovannoni,J.J., Martin,G.B. and Tanksley
,S.D.
TITLE Generation of ESTs from germinating tomato seed
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
FEATURES
source Location/Qualifiers
1..657
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLE11l6"
/tissue_type="whole seedlings"
/dev_stage="7 days post imbibition"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; 7 days post imbibition on water-agar. Mixed stage
whole germinating seedlings from seed coat emergence up
to two centimeters in seeds not showing obvious signs of
germination were discarded."
BASE COUNT 212 a 120 c 147 g 178 t
ORIGIN
Query Match 26.2%; Score 420.2; DB 119; Length 657;
Best Local Similarity 80.7%; Pred. No. 3.le-93;
Matches 517; Conservative 0; Mismatches 118; Indels 6; Gaps 2;
Qy 96 TACAAAGTGAAGCTCTGAAAAAACAACGCGTGAACACCATAAACCACTAGGAAAAA 155
Db 17 TACAAAGTGAAGCTCTGAAAAAACAACGCGTGAACACCATAAACCACTAGGAAAAA 76
Qy 156 AGAGTGCAATTGGTACATTTGATCTCTCTTACCTAGCTTCTTATACACCAAAATT 215
Db 77 AGAATGCAATTAGTACATTTGATTTACCTTACTTACTTATATACCAAAAGTT 136
Qy 216 TCTCATCTATAA---AGGTGCTGAAAACTTTTGACGAGACGCGTGAAAAAATTAAAGATGG 272
Db 137 GATGGTTTACAAAGTTAGGGCTGAAGCTTTGAGAAACAGTAGAANAATTTGAAGATGG 196
Qy 273 ACTGGCCTTAGTATGGTGGATTCTTATCAACTAGCTGGGAAACTTGGAAAGATGAAGA 332
Db 197 TTTGGCTTTAGTGTGGAAGATTTTATCACTTGTCTGGAATAATTGGAAAAAGATGA 256
Qy 333 AGGGGTTTTTCAGGTGGAATACGACGATGACATGATGGTGTACAGGTGACAGTGGCTGT 392
Db 257 AGGAGTATTTAAGTGGAAATATGATGATATGATGATGATGATGATGATGATGATGATG 316

```

```

393 TGCAGAGAGATAGAGAGTTGCGAGATCTTACTG---ATGAGAAGGACCACCAAAATTGCA 449
Db 317 GGCTCAAGAAATTCAGTTGCTAATCTTACTGATCATCATGAAGGATCAATAAATTCGA 376
Qy 450 GGACTTGATTCCTTGTAAATAAAATCTTGAATTTGGAAGGGCTTCATCGCCTCTTTAGC 509
Db 377 AGATTTGATACCTTTATAACAAAATCTTGAATTTAGAAGGGCTTCATAGGCCTCTACTCGC 436
Qy 510 TGTGCAGCTCACCAAGCTCAAGGACGGCTCACCATGGGATTAGCATTTACCATGCTGT 569
Db 437 CGTACAGCTCACCAAGCTTAAAGACGGGCTGGCAATGGATTAGCATTTAATCATGCTGT 496
Qy 570 GCTGGATGGTACTTCGACGTGGCACCTTTATGACCTCGTGGTCCGAGCTTTGCTGTGGGTC 629
Db 497 GCTGGACGGTACTTCACAGTGGCACCTTCATGACTTCTGGCCCGAGCTTTGTAGTGGGGC 556
Qy 630 CACCTCAATTTCTGTCACACCATTCCTTTGAACGAACCAAGGCTCGTAACACTCGAGTCAA 689
Db 557 CACCTCCATCTCGGTCCGACCTTTCTCGAAAGAACTAAAGCTCGTGACACTCGAGTCAA 616
Qy 690 GCTCAACCTCTCTCAACCATCAGATGCACCGCAACATGCTA 730
Db 617 ACTCAACCTATCTAAGCCATCCGACGCGCGGAACATGCCA 657

```

```

RESULT 3
BG126859 649 bp mRNA EST 31-JAN-2001
LOCUS EST472505 tomato shoot/meristem Lycopersicon esculentum cDNA clone
DEFINITION cTOF13L13 5' sequence, mRNA sequence.
ACCESSION BG126859
VERSION BG126859.1 GI:12627047
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 649)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Uterback,T.,
Hansen,C., Ronning,C. and Tanksley,S.
TITLE Generation of ESTs from tomato shoot/meristem tissue
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES
source Location/Qualifiers
1..649
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOF13L13"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."
BASE COUNT 177 a 170 c 148 g 154 t
ORIGIN
Query Match 25.3%; Score 406.4; DB 174; Length 649;
Best Local Similarity 79.3%; Pred. No. 7.8e-90;
Matches 514; Conservative 0; Mismatches 116; Indels 18; Gaps 2;
Qy 385 GTGCTGTTGCAAGAGAGATAGAAGTTGCAGATCTTACTG---ATGAGAAGGACCAACC 441

```

Db 2 GTGGATGAGGCTCAAGAAATTCAGTTGCTAATCTTACTGATCATCATGAAGGATCAAT 61
Qy 442 AAATTCAGAGCTTGATCTCTTGAATAAAATCTTGAATTTGAAGGGCTTCATCGCCCT 501
Db 62 AAATTCAGAGATTGATACCTTATACAAATATCTTGAATTTGAAGGGCTTCATAGGCCCT 121
Qy 502 CTTCTAGCTGTGCAGCTCACCAAGCTCAAGGACGGGCTCACCATTGGATTAGCATTTAAC 561
Db 122 CTACTCGCCGTACAGCTACCAAGCTTAAAGACGGGCTGGCAATGGGATTAGCATTTAAT 181
Qy 562 CATGCTGCTCGTGGAGTACTTCGAGTGGGCACTTTATGACCTCGTGGTCCAGCTTTCG 621
Db 182 CATGCTGCTCGTGGAGTACTTCGAGTGGGCACTTTATGACCTCGTGGTCCAGCTTTGT 241
Qy 622 TGTGGGTCCACCTCAATTTCTGTCACCATCTTGAAGCAAGCAAGCTCGTAAACACT 681
Db 242 AGTGGGCCACCTCCATCTCGTTCGCTCCACCTTTCTCGAAAGAACTTAAGCTCGTACACT 301
Qy 682 CGAGTCAAGCTCAACCTCTCTCAACCATCAGATGCACCGCAACATGCTTAAGTCAGCAAC 741
Db 302 CGAGTCAAACTCAACCTATCTAAGCCATCGGACGCGCGGACATGCCAAGTCGGAACG 361
Qy 742 AAGGTGATGTCCGGCCAACTAGACCACTCTTCGCGGAAAGATATTCAAGTTCTCC 801
Db 362 AAGCGGAGCTTAGCGCCAGCGTGGATCCACCTATGCGTGACAGGGTGTTCAGTTTTC 421
Qy 802 GAGTTAGCAATGACAAATCAAGTCAACAGTCAATGCCAATCAAGGAGGCGCAAT 481
Db 422 GAATCAGCAATGATCAAAATCAAGTCAAAAGTCAATTAACCGAGGAGGCGCAAT 481
Qy 852 -----GACGCCATTTCCACATTCACATCTCCGCGACAGTGTGGCTAGCGCTCACA 906
Db 482 AATACGACGCGTCTCTACTTTCCATCCCTCTCCGCGACAGTGTGGCTAGCTGTACG 541
Qy 907 CGTCCGCGCACTCAAGCCGAGGACTACACTGTGTACACTGTGTGTGTGTGTGTGTGTGT 966
Db 542 CGTCCGAGGACTCTCAAAACCGGAGGAATACAGGTTTACACCGTGTGTGTGTGTGTGT 601
Qy 967 AAAAGGTTGATCTCCATGTCAGAAAGTACTTCGCGCAACCTAAT 1014
Db 602 AAAGAGTCTGATCCGCGGATGCGGAGAGTTACTTCGGAACCTAAT 649

RESULT 4
LOCUS BF051093 642 bp mRNA EST 16-OCT-2000
DEFINITION EST436268 tomato developing/immature green fruit Lycopersicon
esculentum cDNA clone cLEM21F13 5' sequence, mRNA sequence.
ACCESSION BF051093
VERSION BF051093.1 GI:10804989
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 642)
AUTHORS Alcalá, J., Vrábalov, J., White, R., van der Hoeven, R. S., Holt, I. E.,
Liang, F., Hansen, T. S., Craven, M. B., Bowman, C. L., Ronning, C. M.,
Nierman, W., Fraser, C. M., Martin, G. B., Giovannoni, J. J. and Tanksley,
S. D.
TITLE Generation of ESTs from tomato fruit tissue, immature green
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu.
FEATURES
Location/Qualifiers
1. .642

/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEM21F13"
/clone_lib="tomato developing/Immature green fruit"
/tissue_type="fruit"
/dev_stage="immature green (5-35 days post-anthesis)"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmCuadapt; Site_1: EcoRI;
Site_2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and
harvested at 7 day intervals through 35 dpa. Equal masses
of tissue from each stage were combined (including seeds
and locules) prior to mRNA isolation."
BASE COUNT 179 a 128 c 197 g 138 t
ORIGIN
Query Match 25.2%; Score 404.2; DB 143; Length 642;
Best Local Similarity 78.5%; Pred. No. 2.7e-89;
Matches 503; Conservative 0; Mismatches 123; Indels 15; Gaps 1;
Qy 715 GCACCGACATGCTAAGTTCAGCAACCAACGCTGATGTCGCGCCCAACGTAGACCCACCT 774
Db 2 GCGCCGGAACATGCCAAGTCGGAACCAACGCGACGTTAGCGCCGCGGTGATCCACCT 61
Qy 775 CTTGCGGAAGAGTATTCAAGTTCTCCGAGTTAGCAATTTGACAAAATCAAGTCAACAGTC 834
Db 62 ATGCGTGACAGGGTGTTCAGTTTCCGAATCAGCAATTTGATCAAAATCAAGTCAAAAGTC 121
Qy 835 AATGCCAATCAAGGAGA-----GACGCCATTTCCACATTTCCAAATCAACCTC 879
Db 122 AATACTAACCCAGGAGAGAGCGCAATAATACGACGCGGTCTCTACTTTCCAAATCCCTC 181
Qy 880 TCCGCACACGTGTGGCTAGCCGTCAACAGTGCAGCGCCCAACTCAAGCCCGGAGGACTACACT 939
Db 182 TCCGCCACGTGTGGCTAGCTGTACGCGTGGGAGGAGCTCAACCGGAGGAATACAGC 241
Qy 940 GTGTACACTGTGTGTGCTGATTGCAGGAAAAGGTTTGATCTCTCAATGCCAGAAAGTTAC 999
Db 242 GTTACACCGTGTTCGCTGATTTGCCGTAAAGAGTGCATCCGCCGATGCCGAGAGTTAC 301
Qy 1000 TTGCGCAACCTAATTCAGCAATTTTCAGTGACCGCGGAGGTTTGTGTACTAGCAAGC 1059
Db 302 TTGGGAAACCTAATTCAGTCGATTTTCACTGTAACGCGCGGGGATTTGCTGTATCGAAT 361
Qy 1060 CCGATCGAGTTCGCTGCTGGATGATACAAACAGCCATCGTGAAGCATGACGCTAAGGCC 1119
Db 362 CCGATTGAGTTTCCGCGACGGATGATTTCATCAAGCCATAGCGAAGCAGCATGCCAAGGCG 421
Qy 1120 ATTGATGAAAGAAACAGGAGTGGGAGAGCAACCCCAAGATCTTTTCAGTACAAAGATGCT 1179
Db 422 ATTGAAGAAGAACCAAGGAATGGGAGAGTAGTCCCAAGATATTTCCGTACAAAGATGCC 481
Qy 1180 GGAGTGAACCTGTGCTGCTTTGGAAGTTTCGCCAAGTTTCAAGGTTTACGACGTTGATTT 1239
Db 482 AGAGTCAACTGTGTGCGGTTGGGAGCTTCGCCAAGCTTTAAGGAGTATGAGGTTGATTTT 541
Qy 1240 GGATGGGAAAGCCAGAGAGTGTGAGAGTGTTCGAAACAAATAGGTTTGTATGAATGGTG 1299
Db 542 GGATGGGAAAGCCAGAGAGTGTGAGAGTGTTCGAAACAAATAGATTTGACCGGATGGTG 601
Qy 1300 TATTTCTACCAAGGCAAAATATGGAGGAAGCAAGCATTTGATGT 1340
Db 602 TATTTGTATCCAGGCAAAATGGAGGAAGGAATTTGATGT 642
RESULT 5
LOCUS BE919634 588 bp mRNA EST 02-OCT-2000
DEFINITION BE919634 potato leaves and petioles Solanum tuberosum cDNA clone
cSTB2C8 5' sequence, mRNA sequence.
ACCESSION BE919634
VERSION BE919634.1 GI:10445710


```

182 ATCAACTTCTCGAAAATTTGGAAAAGATGATGAAGGAGTATTTAAAGGTGGAATATGATG 241
QY 359 ATGACATGATGGTGTAGAGGTGCACAGTGGCTGTTCACAGAAGATAGAAAGTTGCAGATC 418
Db 242 ATGATATGATGGATGTAGGTTAAATTTGTGGATGAGGCTCAAGAAATTCAGTTGCTTAATC 301
QY 419 TTAAGTCTG---ATGAAGAAGCCACACCAATTCGACGACTTGTATTCCTTGTAAATAAATCT 475
Db 302 TTAAGTCTGATCATGAAGGATCAATAAATTCGAAGATTTGATACCTTATAACAAAATCT 361
QY 476 TGAATTTGAAGGGCTTCATCGCCCTCTTCTAGCTGTGCAGCTCAACAGCTCAAGGACG 535
Db 362 TGAATTTGAAGGGCTTCATAGCCCTCTACTCGCCGTACAGCTCAACAGCTTAAAGACG 421
QY 536 GGCTACCATGGATAGCATTTAACCATGCTGTGTGATGGTACTTCGACGTGGCACT 595
Db 422 GGCTGCAATGGATTAAGCATTTAATCATGCTGTGTGACGGTACTTCCACGTGGCACT 481
QY 596 TTATGACCTCGTGGTCCGAGCTTTGCTGTGGGTCCACCTCAATTTCTGTCCACCAATTC 655
Db 482 TCATGACTTCTTGGGCCAGCTTTGTAGTGGGCCACCTCCATCTCGTCCACCACTTCC 541
QY 656 TTGAACGAACCAAGGCTCGTAACATCGAGTCGAGTCAAGCTCAACCTCTCTCAACCATCAGATG 715
Db 542 TCGAAGAAGAACTAAAGCTCGTACATCGAGTCGAGTCAAACTCAACCTATCTAAAGCCATCGGACG 601
QY 716 CACCCCAACATGC 728
Db 602 CGCCGGAACATGC 614

RESULT 7
BG129408
LOCUS
DEFINITION
EST475054 tomato shoot/meristem Lycopersicon esculentum cDNA clone
CTOF24G12 5' sequence, mRNA sequence.
ACCESSION
BG129408
VERSION
BG129408.1 GI:12629596
KEYWORDS
EST.
SOURCE
tomato.
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE
1 (bases 1 to 651)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
Hansen,C., Ronning,C. and Tanksley,S.
Generation of ESTs from tomato shoot/meristem tissue
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES
Location/Qualifiers
1..651
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CTOF24G12"
/clone.lib="tomato shoot/meristem"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOUR"
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."
203 a 123 c 151 g 174 t

BASE COUNT
ORIGIN
```

```

Query Match 24.1%; Score 386.8; DB 174; Length 651;
Best Local Similarity 79.8%; Pred. No. 5.5e-85;
Matches 519; Conservative 0; Mismatches 122; Indels 9; Gaps 5;

QY 122 CAAACGCTGAACACACATAAACCACTAGGAAAAAAG-AGTGTCAATTTGGTAACATTTGAT 180
Db 1 CAAATCTGAAGCCACAAAACCACTAGGGAAAAAAGAAATGTCAATTAGTAACATTTGAT 60
QY 181 CTTCTCTACTAGCTTTCTATTACAAACCAAAATTTCTCATCTATAA---AGTGTCTGAA 237
Db 61 TTACCTTTACTTACCTTTTATTATTAACCAAAAGTTGATGGTTTACAAATTAGGGGTGAA 120
QY 238 AACTTTGACGAGACGCTGGAAAAAATTAAGATGGAGCTGGCCCTTAGTATTGGTGGATTTTC 297
Db 121 AGCTTTGACGAACACATAGAAAAATTTGAAGATGGTTTGGCTTTTAGTGTGGGAAGATTTT 180
QY 298 TATCAACTAGCTGGGAACTTTGGAAAAAGATGAAGAAGGGTTTTCAGGGTGGAAATACGAC 357
Db 181 TATCAACTTGTCTGAAAAATTTGGAAAAAGATGATGAAGGAGTATTTAAGGTGGAAATATGAT 240
QY 358 GATGACATGATGGTGTAGAGGTGACAGTGGCTGTTCGAGAAGAGATAGAAATTCGAGAT 417
Db 241 GATGATATGGATGGAGTTTGAGGTAATTTGTGGATGAGGCTCAAGAAATTTCAAGTTGCTAAT 300
QY 418 CTTACTG---ATGAAGAAGSCACCAAAATTCGACGACTTGTATTCCTTGTATAAATAATC 474
Db 301 CTTACTGATCATCATGAAGSGATCAATAATTCGAAGATTTGATACCTTATAACAAATC 360
QY 475 TTGAATTTGGAAGGGCTTCATCGCCCTCTTCTAGCTGTGCAGCTCAACAGCTCAAGGAC 534
Db 361 TTGAATTTGAAGGGCTTCATAGGCTCTACTCGCCGTACAGCTCAACAGCTTAAAGAC 420
QY 535 GGCTCACCATGGGATTAGCATTTTAACCATGCTGTGCTGGATGGTACTTCGACGTGGCAC 594
Db 421 GGCTCGGCAATGGGATTAGCATTTTAATCATGCTGTGCTGGACGGTACTTCCACGTGGCAC 480
QY 595 TTTATGACCTCGTGGTCCGAGCTTTTCTGTGGTCCACCTCAATTTCTGTGCCACCATTC 654
Db 481 TTCATGACTTCTCGGCCACGCTTTGTAGTGGGCCACCTCCATCTCGGTCCCACTTTC 540
QY 655 CTTGAACGAACCAAGCTCTCAACACTCGAGTCAAGCTCAACCTCTCTCAACCATCAGAT 714
Db 541 CTCGAAG-ACTAAAGCTCGTGACACTCGAGTCAAACTCAACTATCTAAGCCATCCGAC 599
QY 715 GCACCCGACATGCTAAGTCAGCAACCAACGCTGATGTCGCCGCCAAGCT 764
Db 600 GCGCCGGAACATGCCAAGTC-GGAAACAACGGCGACGTTAGCGCCGACGCT 648

RESULT 8
AW618433
LOCUS
DEFINITION
EST320419 L. pennellii trichome, Cornell University Lycopersicon
pennellii cDNA clone cLPT13E7 5', mRNA sequence.
ACCESSION
AW618433
VERSION
AW618433.1 GI:7324679
KEYWORDS
EST.
SOURCE
Lycopersicon pennellii.
ORGANISM
Lycopersicon pennellii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE
1 (bases 1 to 582)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Lakey,J., Holt,I.E.,
Liang,F., Hansen,T.S., Upton,J., Ronning,C.M., Craven,M.B., Fujii
,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin
,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from wild tomato (Lycopersicon pennellii)
trichomes
Unpublished (1999)
JOURNAL
COMMENT
Contact: David Frisch
```


Db	452	CGTTCAGCTCACCAGCTCAAAAGACGGCTGGCAATGGGATAGCATTAAATCATGCTGT	511
Qy	570	GCTGGATGGTACTTCGACGTGGCCACTTTATGACCTCGTGGTCCGAGCTTTCTGCTGTGGG	627
Db	512	GCTGGATGGTACTCCAGCTGGCACTTCATGACTTCATGGCCCGCCAGCTTTGTAGTGGG	569
RESULT	10		
AW039511			
LOCUS			
DEFINITION		EST281792 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA	
ACCESSION		clone cLET14C2, mRNA sequence.	
VERSION		AW039511	
KEYWORDS		AW039511.1 GI:5898265	
SOURCE		EST.	
ORGANISM		tomato.	
		Lycopersicon esculentum	
		Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots; Asteridae: euasterids I: Solanales; Solanaceae; Solanum; Lycopersicon.	
REFERENCE		1 (bases 1 to 561)	
AUTHORS		D'Ascenzo, M., He, X., Lyman, J., Holt, I. E., Liang, F., Upton, J., Ronning, C. M., Craven, M. B., Fujii, C. Y., Bowman, C. L., Nieman, W., Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley, S. D. and Giovannoni, J.	
TITLE		Generation of ESTs from tomato leaf tissue	
JOURNAL		Unpublished (1999)	
COMMENT		Contact: David Frisch Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrisch@CLEMSON.EDU 5 prime sequence.	
FEATURES		Location/Qualifiers	
source		1..561	
		/organism="Lycopersicon esculentum"	
		/cultivar="Rio Grande Ptor"	
		/db_xref="taxon:4081"	
		/clone="cLET14C2"	
		/clone_1lb="tomato mixed elicitor, BTI"	
		/diverse_type="leaf"	
		/dev_stage="4-6 week old plants"	
		/lab_host="XU1-Blue MRF"	
		/note="vector: pBlueScript SK(-); Site_1: EcoRI; Site_2: XhoI; cLET - Inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisonicotinic acid, BTH, jasmonic acid, ethylene, fenthion, EIX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."	
BASE COUNT		184 a 87 c 129 g 161 t	
ORIGIN			
Query Match		22.18; Score 355.4; DB 110; Length 561;	
Best Local Similarity		80.5%; Pred. No. 3e-77;	
Matches		442; Conservative 0; Mismatches 101; Indels 6; Gaps 2;	
Qy	96	TACAAAGCTGAAGTCTCGTGAACCAACCAACCTGAAACACATACACCTAGGAAAAA	155
Db	13	TACAATGTGAACATCTTGAACAAATCAATGTGAAGCCACAAACCACTAGGAAAAA	72
Qy	156	AGAGTGTCAATTTGGTAACATTTGATCTTCCTTACCTAGCTTCTATTACAAACCAAAAAAT	215
Db	73	AGAATGTCAATTAGTAACATTTGATTACCTTACTTACCTTTTATTATACCAAAAGTT	132
Qy	216	TCTCATCTATAA---AGTGTGTGAACACTTTGCACGACGCGTGGAAAAAATTAAGATGG	272
Db	133	CATGGTTTACAAAGTTGAGGGCTGAAGCTTTGAGGAAACAGCTAGAAAAATTTGAAAGATGG	192

ORIGIN

Query Match 22.1%; Score 355.2; DB 119; Length 544;
Best Local Similarity 80.1%; Pred. No. 3.4e-77;
Matches 436; Conservative 0; Mismatches 93; Indels 15; Gaps 1;

QY 487 GGGCTTCATCGCCCTCTCTTAGCTGTGAGCTCACCAGCTCAAGGACGGGCTCACCATG 546
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GGGCTTCATAGCCCTCTACTCGCGCTCAGAGCTCACCAGCTTAAAGACGGGCTGCAATG 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 547 GGATTAGCATTAACCATGCTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 606
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 GGATTAGCATTAATCATGCTGTGCTGAGCGGTATTCACGCTGGCAGCTTTCATGACTTCC 120
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 607 TGGTCCGAGCTTGTGCTGTGGGTCACCTCAATTTCTGCTCCACCACTTCTTCTTGAACGAACC 666
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 TGGGCCAGCTTGTAGTGGGGCCACCTCCATCTCGGTCCCACTTCTCTGAAAGAACT 180
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 667 AAGGCTCGTAACATCGAGTCAAGCTCAACCTCTCTCAACCATCAGATGACGCCGAACAT 726
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 AAGCTCGTGACATCGAGTCAACCTCAACCTATCTAAGCCATCGACGCGCGGAACAT 240
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 727 GCTAGTCAGCAACCAAGGTGATGTCGGGCCAACGCTAGACCCACCTCTTCTGCGGAAGA 786
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 GCGAAGTCGGAACCAAGCGGACGCTTAGCGCCAGCGTGGATCCACCTATGCGTGACAGG 300
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 787 GTATTCAAGTTCTCCGAGTAGCAATTCACAAATCAAGTCAACAGTCAATGCCAATCA 846
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 GTGTTCAAGTTTCCGAATCAGCAATGATCAAAATCAAGTCAAAAGTCAATACTAACC 360
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 847 GGAGA-----GACGCGATTTCCACATTTCCCAATCACTCTCGGCACACAGTG 891
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 GGAGAAGGCGCAATAATACAGCGCGTCTCTACTTTTCCATCCTCTCCGCGCACGTG 420
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 892 TGGCTAGCGGTACACGCTGCGCGCACTCAAGCCGAGGACTACACTGTGTACACTGTG 951
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 421 TGGCTAGCTGTACACGCTGCGGAGGAGCTCAAAACCGGAGGAATACACGCTTTACACCGTG 480
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 952 TTTGCTGATTGCGAGAAAGGTTGATCCTCAATGCCAGAAAGTTACTTCCGCAACCTA 1011
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 TTCGTTGATGCGGTAAAGAGTGCATCCGCGATGCCGGAGAGTTACTTCCGAAACCTA 540
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1012 ATTC 1015
|||||
Db 541 ATTC 544

RESULT 12

BF053244
LOCUS BF053244 556 bp mRNA EST 16-OCT-2000
DEFINITION EST438474 potato leaves and petioles Solanum tuberosum cDNA clone
CSPB34L5 5' sequence, mRNA sequence.
ACCESSION BF053244
VERSION BF053244.1 GI:10807140
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 556)
van der Hoeven, R.S., Bezzerides, J., Holt, I.E., Liang, F., Cho, J.,
C.M., Fry, W.E., Tanksley, S.D. and Baker, B.

Utterback, T., Hansen, C.L., Doan, B., Bougri, O., Buell, C.R., Ronning

Generation of ESTs from potato leaves and petioles

Unpublished (2000)

Contact: Cathy Ronning

The Institute for Genomic Research

For clone info: please contact Research Genetics, Libraries

Division tel 1-800-711-6195, email cdna@resgen.com.

Location/Qualifiers

1. 556

FEATURES

. source

/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="CSPB34L5"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 185 a 83 c 130 g 158 t
ORIGIN

Query Match 22.0%; Score 353.8; DB 143; Length 556;
Best Local Similarity 82.3%; Pred. No. 7.6e-77;
Matches 432; Conservative 0; Mismatches 87; Indels 6; Gaps 2;

QY 96 TACAAAGTGAAAGTCCTGAAAAAACAACGTAACCAACATAAACCACTAGGAAAAA 155
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 32 TACAAATGTGAACATCTTTGAAAAATCAATGTGAAGCTCAAAAACCACTAGGAAAAA 91
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 156 AGAGTGTCAATTTGGTAACATTTGATCTTCCCTACCTAGCTTTCTATTACACCAAAATTT 215
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92 AGAATGTCAATTTGGTAACATTTGATTTACCTTACTTCTTTTATTATTAACCAAAAGTT 151
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 216 TCTCATCTATAA---AGGTGCTGAAACCTTTGACGAGACGGTGGAAAAAATTTAAAGATGG 272
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 152 GATGTTTTACAAATTTAGGGACTGAAACCTTTGAGGAAACAGTGGAAAAATTTGAAGATGG 211
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 273 ACTGCCCTTAGTATTGGTGGATTTCTATCACTAGCTGGGAACTTGGAAAAGATGAAGA 332
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 212 TTTGGCTTTAGTGTGGGAAGATTTTATCACTTTGGGAAATTTAGGAAAAGATGAAGA 271
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 333 AGGGTTTTTCAGGGTGGAAATACGAGGATGACATGCTGTAGAGGTGACAGTGGCTGT 392
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 272 AGGAGTATTTAAGTGGAAATATGATGATGATGATGATGATGATGATGATGATGATGATG 331
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 393 TGCAGAAGAGATAGAAGTTGCAGATCTTTACTG---ATGAAGAAGGCACCAAAATTTCCA 449
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 332 AGCTGAAGGGATTTGAAGTTGCAAGTCTTACTGATCATCAAGAAGGATTAATAAATTTCCA 391
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 450 GGACTTGATCTCTGTATAAATCTTGAATTTGGAAGGGCTTCATCGCCCTCTCTCTAGC 509
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 392 AGCTTTGATACCTTATAAANAATCTTGAATTTGGAAGGGCTTCATAGGGCTCTCTCTCGC 451
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 510 TGTGAGCTCACCAGCTCAAGGACGGGCTCACCATGGGATTAGCATTTAACCATGCTGT 569
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 452 GCTTACAGCTCACCAGCTCAAGACGGGCTGGCAATGGGATAGCATTTATCATGCTGT 511
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 570 GCTGGATGGTACTTCGACGTGGCAGCTTTATGACCTTCGCTGGTCCGA 614
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 512 GCTGGATGGTACCTCCACGTGGCACTTTCATGACTTCATGCTTCATGGGCCCA 556
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13

BF275133
LOCUS BF275133 786 bp mRNA EST 07-MAR-2001
DEFINITION GA_EB0023D13f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_EB0023D13f, mRNA sequence.
ACCESSION BF275133
VERSION BF275133.1 GI:11206203
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 786)

AUTHORS Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE An integrated analysis of the genetics, development, and evolution
of the cotton fiber
JOURNAL Unpublished (2000)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATCAGCTCATATAGGG
High quality sequence stop: 719.

FEATURES
source
1..786
/organism="Gossypium arboreum"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone_lib="GA_Eb0023D13f"
/tissue_lib="Gossypium arboreum 7-10 dpa fiber library"
/tissue_type="Fibers isolated from bolls harvested 7-10
dpa"
/lab_host="E. coli"
/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 205 a 206 c 202 g 173 t
ORIGIN

Query Match 21.5%; Score 345; DB 146; Length 786;
Best Local Similarity 67.6%; Pred. No. 1.2e-74;
Matches 547; Conservative 0; Mismatches 225; Indels 37; Gaps 3;
QY 546 GGGATTAGCATTTAACCATCGTGTGCTGATGATGCTGACGTGGCACTTTATGACCTC 605
DB 1 GGGGTGCGGATTAACACACGCCATCTTGTGGAAGTCTCCACTTGGCATTTATGAGCTC 60
QY 606 GTGGTCCGAGCTTGTGTGGTCCACCTCAATTTCTGTCGCCACCATCTTGAACGAAC 665
DB 61 ATGGGGCGCAATCTGTAGCGGGTCCAACTCCGCTCAGTCTGCCGTTTCTCGAGCGGAC 120
QY 666 CAAGGCTCGTAACACTCGAGTCAAGCTCAACCTCTCTCAACCATGATGACACCCGGAACA 725
DB 121 CAAAGTCCGAACACCCGCGTGAAGCTGATCTCTCGCTCCCACTCACTCTGTCACAGG 180
QY 726 TGCTAGTCAGCAACCAACGGTGTGTCGGCGCCAGGTAGACCCACCTCTTCGCGCAAG 785
DB 181 TG-----ACGCCAACCAAGGCCCCCGCTGAGGGAGAA 213
QY 786 AGTATTCAAGTTTCCGAGTTAGCAATTGACAAATCAAGTCAACAGTCAATGCCCACTC 845
DB 214 ACTCTTCAGGTTTCTGAGAGTGCCTATCGACAAAGATCAAGTCAAAAGTCAACTCTACCCC 273
QY 846 AGGA-----GAGACGCCCATTTCTCCACATTTCCATCACTCTCGGCACACGCTGTGGT 896
DB 274 ACCATCGACTCTCTTAACCATTTCTCACTTTCCATCTCTAGCTTATCAGCTTTGGCA 333
QY 897 AGCGGTACAGCTGCGCGGCACTCAAGCCGAGGAGTCACTGTGTAGACTGTGTTTC 956
DB 334 CCAGCTAGCCCTAGCAGTGAACCTCAAGCTCAAGATTACACAGTTTTCACCTCTCTCGC 393
QY 957 TGATTCGAGGAAGGGTTGATCCTCAATGCCAGAAAGTTACTTCGGCAACCTATTCA 1016
DB 394 AGATTGCTGTAAGAGGGTTGATCCACCCATGCCAGAAAGTTACTTCGGAACCTGATTCA 453
QY 1017 GGCATTTTTCACAGTCAACGCGCGGCTGTTGTTACTTAAGCCCGATCGAGTTTCGCTGG 1076
DB 454 TGGCATTTTCACGCTCAGCGCAGCGGGTTATTTGTTGGCTATCCACCGCAATTCGGTGC 513
QY 1077 TGGGATGATACAAACGAGTCTGTAGAGCATGACCGTAAAGCCATTGATCAAGAAACAA 1136
DB 514 ATCGTTGGTACAAAAGCTATAGAAGCGCACACCGCAAGGCGCATCGAAGAACGAACAA 573

QY 1137 GGAGTGGGAGCAACCCGAAGATCTTTTCAGTACAAAGATGCTGGAGTGAAGTGTGCTGC 1196
DB 574 GGAGTGGGAGCAGCCCGCAAGATATTTCAGTACAAAGATGCTGGTGTGAAGTGTGCTGC 633
QY 1197 TGTGGAAAGTTCGCCAAGTTTCAAGGTTTACGACGTGGATTTTGGATGGGAAAGCCAGA 1256
DB 634 GGTGGGAGCTCTCCAGGTTTCAAGGTATACGATGTGGATTTTCGGGTGGGAAACCGGA 693
QY 1257 GAGTGTGAGGAGTGTTCGCAACAATAGGTTTTCATGGAATGCTGATTTTCTACCAAGCAA 1316
DB 694 AGGGGTGAGGAGCGGATCCCAACACAGGTTTGTATGGATGGTGTATTTGTATCAAGGAA 753
QY 1317 AAATGAGGAGGAAGCATTTGATCTGGAGA 1345
DB 754 GAGGGGTGGCGGAA-CATTGACGTGGAGA 781

RESULT 14
AW0411172
LOCUS AW0411172 563 bp mRNA EST 18-OCT-1999
DEFINITION EST284036 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
clone cLETL1N19, mRNA sequence.
ACCESSION AW0411172
VERSION AW0411172.1 GI:5899926
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 563)
AUTHORS D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J.,
Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,C.L., Nierman,W.,
Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and Giovannoni
,J.
Generation of ESTs from tomato leaf tissue
Unpublished (1999),
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.

FEATURES
source
Location/Qualifiers
1..563
/organism="Lycopersicon esculentum"
/cultivar="Rio Grande PtoR"
/db_xref="taxon:4081"
/clone_lib="tomato mixed elicitor, BTI"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="X11-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLETL - Inoculated with a variety of disease response
elicitors. Plants exposed to 2,6 dichloroisonicotinic
acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."
BASE COUNT 189 a 84 c 128 g 162 t
ORIGIN

Query Match 21.4%; Score 343.8; DB 110; Length 563;
Best Local Similarity 80.6%; Pred. No. 2.2e-74;
Matches 428; Conservative 0; Mismatches 97; Indels 6; Gaps 2;
QY 96 TACAAAGTGAAGTCTCTGAAAAACAAACCGTGAACACCACTAATACCACTAGGAAAAA 155
||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 31 TACAAATGTGAACATCTTTGAAAAAATCAAAATGTGAAGCCACAAAAACCACTAGGGAATA 90

Qy 156 AGAGTGTCAATTTGGTAACATTTGATCTTCCCTTACTAGCTTTCTATTACAACCAAAATTT 215

Db 91 AGAATGCAATTTAGTAACATTTGATTTACCTTACTTACTGCTTTTATTATACCAAAAGTT 150

Qy 216 TCTCATCTATAA---AGCTGCTGAAAACTTTTACGAGACGCTGGAAAAAATTTAAAGATGG 272

Db 151 GATGGTTTACAAGTTAGGGGCTGAAAGCTTTTGAGGAACACGTAGAAAAAATTTGAAAGATGG 210

Qy 273 ACTGGCCTTAGTATTGGTGGATTTCTATCACTAGCTGGGAACTTGGAAAAAGATGAAGA 332

Db 211 TTGCTTTTAGTGTGGGAAGATTTTATCAACTTCTGGAAAAATTTGGAAAAAGATGATGA 270

Qy 333 AGGGTTTTTCAGGTGGGAATACGACGATGACATGGTGTAGAGTGACAGTGGCTGT 392

Db 271 AGGAGTATTAAAGTGGGAATATGATGATATGATGGAGTTGAGGTAAATTTGGATGA 330

Qy 393 TGCAGAAGAGATAGAAGTTGCGAGATCTTACTGATGAA---GAAGGCCACCAAAATTTGCA 449

Db 331 GGCTCAAGAAATTCAGTTGCTAATCTTACTGATCATCATGAAGGATCAATAAATTCCA 390

Qy 450 GGACTTGATCTCTGTATATAAATTTGAATTTGAAGGGGCTTATCGCCCTCTTCTAGC 509

Db 391 AGATTGTATACCTTATACAAAAATCTTGAATTTTGAAGGGGCTTATAGGCGCTTACTCGC 450

Qy 510 TGTGAGCTCACCAAGCTCAAGGAGGGGCTCACCATGGGATTTAGCATTTTAACCATGCTGT 569

Db 451 CGTACAGCTCACCAAGCTTAAAGAGGGGCTGGCAATGGGATTAGCATTTTAATCATGCTGT 510

Qy 570 GCTGGATGCTACTTGGAGTGGGCACTTTATGACCTCGCTGTCGCGAGCTTTG 620

Db 511 GCTGGAGCGTACTTCCAGTGGCACTTTCATGACTTCTCTGGGCCCGAGCTTTG 561

RESULT 15

AW617090 545 bp mRNA EST 24-MAR-2000

LOCUS EST323501 L. hirsutum trichome, Cornell University Lycopersicon

DEFINITION hirsutum cDNA clone cLHT21F17 5', mRNA sequence.

ACCESSION AW617090

VERSION AW617090.1 GI:7323200

KEYWORDS EST.

SOURCE Lycopersicon hirsutum.

ORGANISM Lycopersicon hirsutum

Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE 1 (bases 1 to 545)

AUTHORS van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.

TITLE Generation of ESTs from wild tomato (Lycopersicon hirsutum) trichomes

JOURNAL Unpublished (2000)

COMMENT Contact: David Frisch

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 4366

Fax: 864 656 4293

Email: dfrisch@CLEMSON.EDU

5 prime sequence.

Location/Qualifiers

1..545

/organism="Lycopersicon hirsutum"

/db_xref="taxon:62890"

/clone="cLHT21F17"

/clone_lib="L. hirsutum trichome, Cornell University"

/tissue_type="trichome"

/dev_stage="mixed stages"

/note="Leaves of various stages were shaken in liquid

nitrogen, shearing off trichomes. This procedure yielded a mixture of cells which is highly enriched for trichome cells, likely with minor contaminations of other types of leaf cells."

BASE COUNT 133 a 153 c 135 g 124 t

ORIGIN

Query Match 20.7%; Score 332.2; DB 118; Length 545;

Best Local Similarity 77.4%; Pred. No. 1.6e-71;

Matches 422; Conservative 0; Mismatches 108; Indels 15; Gaps 1;

Qy 565 GCTGTGCTGGATGTACTTTCGACGTGGCAGCTTATATGACCTCTGCTGCTCGAGCTTTGCTGT 624

Db 1 GCTGTGCTGGAGCGGTACTTTCACGTGGCAGCTTATGACTTCTCTGGGCCAGCTTTGTAGT 60

Qy 525 GGGTCCACCTCAATTTCTGTCCCAACCATTTCTTGAAGCAACCAAGGCTCGTAACACTCGA 684

Db 61 GGGGCCACCTCCATCTCGGTCCCACTTTCTCTGACCGCACTAAAGCTCGTGACACTCGA 120

Qy 685 GTCAAGCTCAACCTCTCTCAACCATGAGATGCACCCGAACATGCTAAAGTCAGCAACCAAC 744

Db 121 GTCAAACTCAACCTTATCTAAGCCATCCGACGACCGCAACATGCCAAGTTCGGAACGAAAC 180

Qy 745 GGTGATGTCCCGGCCAACGTAGACCCACCTCTTCGCGAAAGAGATTTCAAGTTCTCCGAG 804

Db 181 GGCACATTTAGCGCCAGCGTGGATCCACCTATGCTGACAGGGTGTTCAGTTTCCGAA 240

Qy 805 TTACGAATTTGACAAATCAAGTCAACAGTCAATGCCAACTCAGGAGA----- 851

Db 241 TCAGCGATTTGATCAAAATCAAGTCAAAAGTCAATACGAATCCACCAGAGGACAGATAAC 300

Qy 852 --GAGCGCATTTCTCCACATTCCTCAATCACTCTCCGACACAGTGTGCTAGCGTCACAGT 909

Db 301 AGTCGCTGCTTCTTCTTCTTCCAAATCCTCTCCCGGACGCTGCTAGCTGTGCTAGCTGT 360

Qy 910 GCGGCCCAACTCAAGCCCGAGGACTACACTGTGTACACTGTGTGTTTGTGTTGCTAGTTCAGGAAA 969

Db 361 GCGAGGCGAGCTGAAACCCGAGGAGATACAGGTTTACACCGTGTTCGCTGATTCGCGTAAA 420

Qy 970 AGGTTTGTATCTTCCAAATGCCAGAAAGTTACTTTCGGCAACCTTAATTCAGGCAATTTTCA 1029

Db 421 AGAGTCGATCCGCGGATGCGGAGAGTTACTTCGGAACCTTAATTCAGGCGATTTTCACT 480

Qy 1030 GTGACCCGCGGAGCTTTGTTTACTAGCAAGCCCGATCGAGTTTCGCTGCTGAGTATACAA 1089

Db 481 GTAACGCTCGGGATTTGCTGCTATCGAATCCGATTGAGTTTGGCGGAGGATGATTTCAT 540

Qy 1090 CAAGC 1094

Db 541 GAAGC 545

Search completed: November 5, 2001, 18:00:51

Job time: 14891 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 16:46:27 ; Search time 8904.87 Seconds
(without alignments)
2569.021 Million cell updates/sec

Title: US-08-894-356C-4
Perfect score: 1479
Sequence: 1 CCGTGATCGAACGCTGTAGA.....TTGAAAAAAAAAAAAAAAAAAAA 1479

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: gb_ba1:*
 - 2: gb_ba2:*
 - 3: gb_ba3:*
 - 4: gb_in1:*
 - 5: gb_in2:*
 - 6: gb_in3:*
 - 7: gb_om:*
 - 8: gb_ov:*
 - 9: gb_pat1:*
 - 10: gb_pat2:*
 - 11: gb_ph:*
 - 12: gb_pl1:*
 - 13: gb_pl2:*
 - 14: gb_pl3:*
 - 15: gb_pl4:*
 - 16: em_ba1:*
 - 17: em_ba2:*
 - 18: em_fun:*
 - 19: em_htgo_hum:*
 - 20: em_htgo_inv:*
 - 21: em_htgo_rod:*
 - 22: em_htg_hum1:*
 - 23: em_htg_hum2:*
 - 24: em_htg_hum3:*
 - 25: em_htg_hum4:*
 - 26: em_htg_hum5:*
 - 27: em_htg_hum6:*
 - 28: em_htg_hum7:*
 - 29: em_htg_hum8:*
 - 30: em_htg_inv1:*
 - 31: em_htg_inv2:*
 - 32: em_htg_other:*
 - 33: em_htg_rod:*
 - 34: em_hum1:*
 - 35: em_hum2:*
 - 36: em_hum3:*
 - 37: em_hum4:*
 - 38: em_hum5:*
 - 39: em_hum6:*
 - 40: em_hum7:*
 - 41: em_in:*
 - 42: em_om:*
 - 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vi:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_v1:*
- 59: gb_v12:*
- 60: gb_htg1:*
- 61: gb_htg2:*
- 62: gb_htg3:*
- 63: gb_htg4:*
- 64: gb_htg5:*
- 65: gb_htg6:*
- 66: gb_htg7:*
- 67: gb_htg8:*
- 68: gb_htg9:*
- 69: gb_htg10:*
- 70: gb_htg11:*
- 71: gb_htg12:*
- 72: gb_htg13:*
- 73: gb_htg14:*
- 74: gb_htg15:*
- 75: gb_htg16:*
- 76: gb_htg17:*
- 77: gb_htg18:*
- 78: gb_htg19:*
- 79: gb_htg20:*
- 80: gb_htg21:*
- 81: gb_htg22:*
- 82: gb_htg23:*
- 83: gb_htg24:*
- 84: gb_htg25:*
- 85: gb_pr1:*
- 86: gb_pr2:*
- 87: gb_pr3:*
- 88: gb_pr4:*
- 89: gb_pr5:*
- 90: gb_pr6:*
- 91: gb_pr7:*
- 92: gb_pr8:*
- 93: gb_pr9:*
- 94: gb_ro1:*
- 95: gb_ro2:*
- 96: gb_in4:*
- 97: gb_pr10:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1478.6	100.0	1479	10	E12756
2	1447.6	97.9	1476	12	AB029340
3	218.4	14.8	1622	10	E12754
4	218.4	14.8	1622	12	AB026494
5	190.6	12.9	1508	10	E12757
6	172.2	11.6	1679	12	AB010708
7	172.2	11.6	1703	10	E12753
8	110.8	7.5	66237	12	AB016892

- E12756 Perilla oci
- AB029340 Perilla f
- E12754 Gentianatri
- AB026494 Gentiana
- E12757 Senecio cru
- AB010708 Gentiana
- E12753 Gentianatri
- AB016892 Arabidops

9	110.8	7.5	101176	12	AC002560	AC002560 Genomic s
10	99.6	6.7	119914	12	AC003027	AC003027 Arabidops
11	76.8	5.2	1648	9	AX025514	Sequence
12	75.4	5.1	85690	12	AB028618	Arabidops
13	75.4	5.1	136047	74	AC069470	Arabidops
c 14	68.8	4.7	136047	74	AC069470	Arabidops
c 15	65.8	4.4	7218	10	I66494	Sequence 14
c 16	63.6	4.3	85690	12	AB028618	Arabidops
c 17	63.4	4.3	82360	13	AP000606	Arabidops
c 18	63.4	4.3	341064	75	AC074226	Arabidops
19	57.2	3.9	7218	10	I66494	Sequence 14
c 20	57	3.9	78379	12	AB006696	Arabidops
c 21	56.8	3.8	71623	77	AC087724	Arabidops
22	52.6	3.6	111338	13	ATF16L2	Arabidops
c 23	52	3.5	93127	74	AC069373	Homo sapi
c 24	52	3.5	186857	74	AC069336	Homo sapi
c 25	51	3.4	14553	4	AE001391	Plasmodiu
c 26	50.8	3.4	46739	13	AP002052	Arabidops
c 27	50.2	3.4	36075	6	CELT27A1	Caenorhab
c 28	49.4	3.3	253305	96	PFMAL3P7	Plasmodiu
c 29	49	3.3	167336	66	AC021350	Homo sapi
c 30	48.4	3.3	156060	60	AC004153	Plasmodiu
31	47.8	3.2	132781	62	AC011689	Homo sapi
32	47.8	3.2	144014	62	AC011685	Homo sapi
33	47.8	3.2	179060	75	AC074382	Homo sapi
c 34	47.2	3.2	42879	93	HSAC000123	Human Cos
c 35	47.2	3.2	45508	93	HSAC000110	Human Cos
c 36	47.2	3.2	245802	60	AC066279	Plasmodiu
c 37	47	3.2	759	53	CNS06QXV	T7 end of
c 38	47	3.2	12029	4	AE001430	Plasmodiu
c 39	47	3.2	63954	77	AC087789	Homo sapi
c 40	47	3.2	93196	67	AC022085	Homo sapi
41	47	3.2	169546	60	AC004157	Plasmodiu
c 42	47	3.2	180301	65	AC018362	Homo sapi
c 43	46.8	3.2	2437	5	AF006380	Tetrahyme
c 44	46.8	3.2	194264	90	AL356956	Human DNA
45	46.4	3.1	2452	5	AF337815	Dictyoste

ALIGNMENTS

RESULT	1
E12756	
LOCUS	E12756 1479 bp DNA PAT 24-JUN-1998
DEFINITION	Perilla oclimoides mRNA for acyltransferase, partial cds.
ACCESSION	E12756
VERSION	E12756.1 GI:3251588
KEYWORDS	JP 1997070290-A/4.
SOURCE	unidentified.
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 1479)
AUTHORS	Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y., Yonekura,K., Mizutani,M. and Kusumi,T. .
TITLE	GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL	Patent: JP 1997070290-A 4 18-MAR-1997; SUNTOBY LTD
COMMENT	OS Perilla oclimoides PN JP 1997070290-A/4 PD 18-MAR-1997 PF 30-JAN-1996 JP 1996046534 PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHIRO, PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAARI PC C12N15/09,A01H1/00,C07K14/42,C12N9/10,(C12N9/10, PC C12R1:865), PC (C12N9/10,C12R1:19); CC strandedness: Double; CC topology: Linear; FH Key Location/Qualifiers

FT	source	1. .11479	/organism='Perilla oclimoides' FT
FT	/tissue_type='leaves'		
FT	CDS	1. .11479	/clone='pSAT208'
FT	Location/Qualifiers		
FEATURES	1. .11479	/organism='unidentified'	
BASE COUNT	420 a	316 c	331 g 411 t 1 others
ORIGIN			
Query Match 100.0%; Score 1478.6; DB 10; Length 1479;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 1478; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CCGTGATCGAAACGTGTAGAGTTGGCGCGCGCGGACTCGTGGCGGAGCAATCGGTGC	60
Db	1	CCGTGATCGAAACGTGTAGAGTTGGCGCGCGCGGACTCGTGGCGGAGCAATCGGTGC	60
Qy	61	CGCTCACATTTCTGCACATGACGTGGCTGCATTTTCATCCCATGCTTCAGCTCTCTTCT	120
Db	61	CGCTCACATTTCTGCACATGACGTGGCTGCATTTTCATCCCATGCTTCAGCTCTCTTCT	120
Qy	121	ACGAATTCCTTTGCCAAGCAACATTTTTCAGAAATCCATCGTTCCAAACTCAACAAT	180
Db	121	ACGAATTCCTTTGCCAAGCAACATTTTTCAGAAATCCATCGTTCCAAACTCAACAAT	180
Qy	181	CTCTCTCTAAACTCTCATATACATTTCTCCCTCTCTCATGCAATTTAATCTACCTTCAT	240
Db	181	CTCTCTCTAAACTCTCATATACATTTCTCCCTCTCTCATGCAATTTAATCTACCTTCAT	240
Qy	241	CCCGGAGAAAATCGCGAGTTTCGGTATCTATCCGGGAGCTCGGTTCTTTTCACCATCG	300
Db	241	CCCGGAGAAAATCGCGAGTTTCGGTATCTATCCGGGAGCTCGGTTCTTTTCACCATCG	300
Qy	301	CAGATCTAGCGAGACATTCGATGATCTCGTCGGAATCGTCAGAAATCCCGGTAGGC	360
Db	301	CAGATCTAGCGAGACATTCGATGATCTCGTCGGAATCGTCAGAAATCCCGGTAGGC	360
Qy	361	TCTACAACTTTTGCCTAAATTTGCCGCCATTTGTGGAAGAATCCGATAGAAACTCTTCC	420
Db	361	TCTACAACTTTTGCCTAAATTTGCCGCCATTTGTGGAAGAATCCGATAGAAACTCTTCC	420
Qy	421	AAGTTTCGCGGTGAGGTGACTCTTTTCCAGCGCGAGGCGTGGTATTGGAATAGCAA	480
Db	421	AAGTTTCGCGGTGAGGTGACTCTTTTCCAGCGCGAGGCGTGGTATTGGAATAGCAA	480
Qy	481	CGCATCACACCGTTAGCGACGCCGCTGTTTCTCGCGCTTTATACGGCTTGGTCTTCAA	540
Db	481	CGCATCACACCGTTAGCGACGCCGCTGTTTCTCGCGCTTTATACGGCTTGGTCTTCAA	540
Qy	541	TGAGCAACACATTTGAAATGAAGATGAAGATGAAGATTTTAAATCTTTGGCAGTTTCG	600
Db	541	TGAGCAACACATTTGAAATGAAGATGAAGATGAAGATTTTAAATCTTTGGCAGTTTCG	600
Qy	601	ATAGATCGGTCAATAATATCCGACGAAATTTGACTTCATTTTGGAGAAACGGCTTAA	660
Db	601	ATAGATCGGTCAATAATATCCGACGAAATTTGACTTCATTTTGGAGAAACGGCTTAA	660
Qy	661	AATTTCCCTTCGAATCTCGTCAATCCCTCATATCCGACGCGGATTCGAAACCGGTTCC	720
Db	661	AATTTCCCTTCGAATCTCGTCAATCCCTCATATCCGACGCGGATTCGAAACCGGTTCC	720
Qy	721	TTTTTACCCCAATCCAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT	780
Db	721	TTTTTACCCCAATCCAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT	780
Qy	781	TAGTCCATCTCTCATCTTTTGTAGGATTTAGGATTTAGGATTTAGGATTTAGGATTTAG	840
Db	781	TAGTCCATCTCTCATCTTTTGTAGGATTTAGGATTTAGGATTTAGGATTTAGGATTTAG	840

FEATURES		Location/Qualifiers	
Source		1..1476	
		/organism="Perilla frutescens"	
		/db_xref="taxon:48386"	
		/tissue_type="leaf"	
CDS		<1..1:43	
		/codon_start=3	
		/product="anthocyanin acyltransferase"	
		/protein_id="BA93475.1"	
		/translation="VIETCRVGGPPDPSVAEQSVPLTFDFDWTWLFHFMQLQLFLYEPFC	
		SKHFSESVNPKLQKSLSLTLHFFPLSLNLIYVSSPERKMPFERYLSGDSVSFTIAES	
		SDDFDLVGNRPESVRLNFKLPPIVVEESDRKLFQVFAVOVTLFPGRGVGIGIAT	
		HITVSDAPSFLAFITAWSSMSKHIEDDEEFKSLPFDVRSVIKYPTFDLSIYWRNA	
		LKFLQSRHPSLPTDRIITFTVFTQSKIKKLKGIQSRVPSLVHLSFVAIAAYMWAG	
		ITKSFADQDQDNEDAFFLIPVDLRPLDPPVPENFTFGNCLSLTALPRMRRLVLVBKG	
		VFLAEVIAAEIKKRINDILETVEKSPKIRKALOKSYFVSGVSKLIDLVLGDFGW	
		GKARKQEILSIDGEKYAMTLCKARDFEGGLEVCLSLPKDKMDAFAYFSLGNG"	
BASE COUNT		419 a 315 c 331 g 411 t	
ORIGIN			
Query Match		97.9%; Score 1447.6; DB 12; Length 1476;	
Best Local Similarity		99.6%; Pred. No. 0;	
Matches 1472; Conservative		0; Mismatches 4; Indels 2; Gaps 2;	
QY	1	CCGTGATCGAAAGCTGTAGAGTTGGGCCGCCGCCGACTCGGTGGCGGAGCAATCGGTGC	60
Db	1	CCGTGATCGAAAGCTGTAGAGTTGGGCCGCCGCCGACTCGGTGGCGGAGCAATCGGTGC	60
QY	61	CGCTCACATTTCTTCGACATGACGTGGCTGCAATTTTTCATCCCATGCTTCAGCTCCTCTTCT	120
Db	61	CGCTCACATTTCTTCGACATGACGTGGCTGCAATTTTTCATCCCATGCTTCAGCTCCTCTTCT	120
QY	121	ACGAATTCCTTGTTCGACGCAACATTTTTCAGAAATCCATCGTTCCAAAACCTCAACAAT	180
Db	121	ACGAATTCCTTGTTCGACGCAACATTTTTCAGAAATCCATCGTTCCAAAACCTCAACAAT	180
QY	181	CTCTCTCTAAACTCTCATACACTTCTTCCCTCTCTCATGCAATTTAATCTACCCCTTCAT	240
Db	181	CTCTCTCTAAACTCTCATACACTTCTTCCCTCTCTCATGCAATTTAATCTACCCCTTCAT	240
QY	241	CCCGGAGAAATCGCGGAGTTTCGGGTATCTATCCGGGGACTCGGTTCCTTTCCACCATCG	300
Db	241	CCCGGAGAAATCGCGGAGTTTCGGGTATCTATCCGGGGACTCGGTTCCTTTCCACCATCG	300
QY	301	CAGAAATCTAGCGAGCTTCGATGATCTCGTGGAAATCGTCGAGAAATCTCCCGTTAGGC	360
Db	301	CAGAAATCTAGCGAGCTTCGATGATCTCGTGGAAATCGTCGAGAAATCTCCCGTTAGGC	360
QY	361	TCTACACTTTGTCCCTAAATTCGCCCCCATTCGCGAAGATCCGATAGAAATCTCTTCC	420
Db	361	TCTACACTTTGTCCCTAAATTCGCCCCCATTCGCGAAGATCCGATAGAAATCTCTTCC	420
QY	421	AAGTTTTCGCGCGTCGAGGTGACTCTTTTCCAGGCCCGAGGCGTGGTATTTGGAATAGCAA	480
Db	421	AAGTTTTCGCGCGTCGAGGTGACTCTTTTCCAGGCCCGAGGCGTGGTATTTGGAATAGCAA	480
QY	481	CGCATCACACCGTTAGGAGCGCCCGCGTGTTCCTCGGGTTTATAACGGCTTGGTCTTCAA	540
Db	481	CGCATCACACCGTTAGGAGCGCCCGCGTGTTCCTCGGGTTTATAACGGCTTGGTCTTCAA	540
QY	541	TGAGCAACACATTGAAATGAAGATCAAGATGAAGAAATTTAAATCTTTGCCAGTTTTCG	600
Db	541	TGAGCAACACATTGAAATGAAGATGAAGAAATTTAAATCTTTGCCAGTTTTCG	600
QY	601	ATAGATCCGTATATAATATCCGACGAAATTTGACTCCATTTATTGGAGAAACCGCTAA	660
Db	601	ATAGATCCGTATATAATATCCGACGAAATTTGACTCCATTTATTGGAGAAACCGCTAA	660
QY	661	AATTTCTTTGCAATCTCGTCAATCCCTCATTCACCGACCGACCGCATTCGAAACCGTTCG	720
Db	661	AATTTCTTTGCAATCTCGTCAATCCCTCATTCACCGACCGACCGCATTCGAAACCGTTCG	720
RESULT 2			
AB029340			
LOCUS		1476 bp mRNA PLN	04-APR-2000
DEFINITION		Perilla frutescens mRNA for anthocyanin acyltransferase, partial cds.	
ACCESSION		AB029340	
VERSION		AB029340.1 GI:7415645	
KEYWORDS		anthocyanin acyltransferase.	
SOURCE		Perilla frutescens leaf cDNA to mRNA.	
ORGANISM		Perilla frutescens	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		Asteridae; eusteroles I; Lamiales; Lamiaceae; Perilla.	
REFERENCE			
AUTHORS		1 (sites)	
		Sakakibara,K.Y., Tanaka,Y., Mizutani,M.F., Fujiwara,H., Fukui,Y.,	
		Ashikari,T., Yamauchi,M. and Kusumi,T.	
TITLE		Molecular and biochemical characterization of a novel	
		hydroxycinnamoyl-CoA: anthocyanin 3-O-glucoside-6	
JOURNAL		Unpublished (1999)	
REFERENCE		2 (bases 1 to 1476)	
AUTHORS		Sakakibara,K.Y. and Tanaka,Y.	
TITLE		Direct Submission	
JOURNAL		Submitted (24-JUN-1999) to the DDBJ/EMBL/GenBank databases. Keiko Y	
		Sakakibara, Suntory Research Center, Fundamental Research, Plant	
		Biotech; Shimamoto-cho, Wakayamada, 1-1-1, Mishima-gun, Osaka	
		618-8503, Japan (E-mail:keiko_sakakibara@suntory.co.jp,	
		Tel:+81-75-962-8807, Fax:+81-75-962-8262)	

Qy 721 TTTTCACCAATCCAAATTAAGAAATTAAGAGGTTGGATTGATTCAGTCCAGAGTTCCAAGTT 780
Db 721 TTTTCACCAATCCAAATTAAGAAATTAAGAGGTTGGATTGATTCAGTCCAGAGTTCCAAGTT 780
Qy 781 TAGTCACATCTCANTCTTTTGTAGCGATTGACAGCTTATATGTGGCTGGCATAACGAAT 840
Db 781 TAGTCACATCTCANTCTTTTGTAGCGATTGACAGCTTATATGTGGCTGGCATAACGAAT 840
Qy 841 CATTACAGCAGATGAAGACCAAGACACAGGAGTGCATTTTCTTGATTCGGTCCGATC 900
Db 841 CATTACAGCAGATGAAGACCAAGACACAGGAGTGCATTTTCTTGATTCGGTCCGATC 900
Qy 901 TAAGCCACGATTAGATCCGCGGTTCTGAAATTAATCTCGGAACGCTTATCGTAGC 960
Db 901 TAAGCCACGATTAGATCCGCGGTTCTGAAATTAATCTCGGAACGCTTATCGTAGC 960
Qy 961 CGCTCCGAGAAATCGGCGGAGAGCTGGTGGGAGAAAGGGTGTCTTGGCAGCTG 1020
Db 961 CGCTCCGAGAAATCGGCGGAGAGCTGGTGGGAGAAAGGGTGTCTTGGCAGCTG 1020
Qy 1021 AGTAAATCGCGCGAGATATAAAGGATCAACGACAAAGAGATATTAGAAACGGTGG 1080
Db 1021 AGTAAATCGCGCGAGATATAAAGGATCAACGACAAAGAGATATTAGAAACGGTGG 1080
Qy 1081 AGAAATGTCGCGCGAGATTTCGTAAGCGTTGCAGAAATCATATTTTCGTTGCGAGGAT 1140
Db 1081 AGAAATGTCGCGCGAGATTTCGTAAGCGTTGCAGAAATCATATTTTCGTTGCGAGGAT 1140
Qy 1141 CGAGCAAGCTAGATCTTACGCTGAGATTTTGGATGGGGAAGCGAGAAAGCAAGAA 1200
Db 1141 CGAGCAAGCTAGATCTTACGCTGAGATTTTGGATGGGGAAGCGAGAAAGCAAGAA 1200
Qy 1201 TATTGTCGATTGATGGGAGAAATATGCAATGACGCTTTGTAAAGCCAGGATTTGCGAAG 1260
Db 1201 TATTGTCGATTGATGGGAGAAATATGCAATGACGCTTTGTAAAGCCAGGATTTGCGAAG 1260
Qy 1261 GAGGATTGAGGTTTGTGCTTTTSCCTAAGGACAAATATGATGATGATGATGATGATGAT 1320
Db 1261 GAGGATTGAGGTTTGTGCTTTTSCCTAAGGACAAATATGATGATGATGATGATGATGAT 1320
Qy 1321 TTTCACTGGGAATTAATGTTAAATTAATGATGATGATGATGATGATGATGATGATGATGAT 1380
Db 1321 TTTCACTGGGAATTAATGTTAAATTAATGATGATGATGATGATGATGATGATGATGATGAT 1380
Qy 1381 ATTAATTAAGTCTTCAGTAAGTGAAGAAATATCCCTATTATATATATATATATATATATAT 1440
Db 1381 ATTAATTAAGTCTTCAGTAAGTGAAGAAATATCCCTATTATATATATATATATATATATAT 1440
Qy 1441 AAATAAGTGAAGGCTCTTTGAAAAAAGGAAAAA 1478
Db 1440 AAATAAGT-TAAGGCTCTTTGAAAAAAGGAAAAA 1476

RESULT 3
E12754 1622 bp DNA PAT 24-JUN-1998
LOCUS Gentianatriflora mRNA for acyltransferase, complete cds.
DEFINITION
ACCESSION E12754
VERSION E12754.1 GI:3251586
KEYWORDS JP 1997070290-A/2.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1622)
AUTHORS Ashikari.T., Tanaka.Y., Fujiwara.H., Nakao.M., Fukui.Y.,
Yonekura.K., Mizutani.M. and Kusumi.T..
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 2 18-MAR-1997;
SUNTORY LTD
COMMENT OS Gentianatriflora
PN JP 1997070290-A/2
PD 18-MAR-1997

PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO,
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAAKI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
C12R1:865),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1622
FT /organism="Gentiana triflora" FT
FT /variety="japonica",
FT /tissue_type="petal",
FT /clone="pgAT106"
FT CDS 35..1474
FT /product="acyltransferase".
FEATURES
source
1..1622
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 478 a 345 c 322 g 477 t
ORIGIN
Query Match 14.8%; Score 218.4; DB 10; Length 1622;
Best Local Similarity 50.8%; Pred. No. 9.5e-43;
Matches 742; Conservative 0; Mismatches 661; Indels 57; Gaps 7;
Qy 3 GTGATCGAAAGCTGTAGAGTTGGCGCGCGGACTCGGTGGCGGAGCAATCGGTGGCG 62
Db 62 GTTCTTGAGAAATCGCGTGTGGCGCCACCACCGGACCGCTGCGCGAGTTTACAGTCCA 121
Qy 63 CTCACATCTTCGACATGACGTGGCTGCATTTTCATCCCATGCTTCAGCTCCTCTTCAC 122
Db 122 CTGCTGTTTTTCGACATGCGATGGTTGATCTCTGATGAGAACACCATCTGCTGCTAC 181
Qy 123 GAATTCCTCTGTCCCAAGCAACATTTTTCAGAAATCCATCGTTCACAAACTCAACAATCT 182
Db 182 AGATTCCCGCATCTCTGTGTCACAC---TCTAAATTTTCAATTCATCCATTAATTCGTC 238
Qy 183 CTCTCTAAACTCTCATACACTTCTTCCCTCTCTCATGCAATTTAATACCTTTCA--- 239
Db 239 CTTTCCCTTGTCTCAACACACTTCTTCCGTTAGCGGGAATTTGATTTGGCGGTAGAT 298
Qy 240 TCCCGGAGAAATGCCGGAGTTTCGGTATCTATCCGGGGACTCGGTTTCTTTCCACCATC 299
Db 299 TCCTCCGATAGAAATGCCGGAGTTGCGTTTACAAGAAAGGGGACTCGGTTTCTTTTAAACAAT 358
Qy 300 GCAGAAATCTAGCGGACTTCGATGATCTCGTTCGGAATTCGTCCAGAAATCTCCCTTAGG 359
Db 359 GCAGAAATCGAGATGGATTTGATTAATCTCGCGGAGATCATCAGAGGATCTTATATA 418
Qy 360 CTCTACAACTTTGTCCTTAAATTTGCCGCCATTTGTCGAAGAAATCCGATAGAAATCTTC 419
Db 419 TTAACAGATTTGATTCGCCAGCTGCCAGAACGATTTGTAACCTCCGCGGACGAAGTATTA 478
Qy 420 CAAGTTTTTCGGGTCAGGTGACTCTTTTCCAGGCCGAGGGCTCGGTATTTGGAATAGCA 479
Db 479 CCACTTTTTGTGTTTACAGGTGACGGTGTCTTCCAAACACCGGTATATGATTTGGACGAAT 538
Qy 480 ACGCATCACCGTTAGCGAGCGCCCGCTCGTTTCTCGCGTTTATAACGGCTTGTCTTCA 539
Db 539 CTTTATCAAGTCTTGGTGATGCCAGTTCTTTCTGCGATTTTATAAATATATGGGTTTG 598
Qy 540 ATGACCAACACATTTGAAAATGAAGATGAAGATGAAGAAATTTAAATCTTTGCCAGTTTC 599
Db 599 GTTGACAAATCCAATGGAGATTCATTAAAGTTCCCTTCCACTTTCTCTCTACCTATGAC 658
Qy 600 GATAGATCCGTCATAAAATATATCCGACGAATTT-----GACTCCATTTATTTGGAGAAC 653
Db 600 GATAGATCCGTCATAAAATATATCCGACGAATTT-----GACTCCATTTATTTGGAGAAC 653

Db	659	GACGATCTGTGGTGAAGATCCATTTTCATATTCGTCGAAATCTACAAATGAAGAAAA	718
Qy	654	GGCTAAATATTCCTTTGCAATCTGTC-----ATCCCTCATATTACCGAGCGACCCCAT	707
Db	719	CTGCTCAATCTCAGGGCACACTACTGTTCTAAATCCAGCAATTTCTTAAAGATGAAGTT	778
Qy	708	CGAACACCGTCTGTTTTCACCAATCCAAATTAAGAAATTAAGAGGTTGGATTTCAGTCC	767
Db	779	CGAGCCACCTTCATCCTACACCTATTGATATCATCAAGCTCAAGAAATTCATTCGTCA	838
Qy	768	AGAGTTCGAAGTTTACTC-----CATCTCTCATCTTTTGTAGGATTCGA	812
Db	839	AAAAATCGCAACTTACCCGGTAGTAGTAATAATCTGTCACTTTCACGGTGCATCT	898
Qy	813	GCTTATATGTGGCTGGCATACGAATCATTC-----ACACGAGATGAAGAC	860
Db	899	GCNCTGATCTGGACATCTTGTGGAATCATTTAGACACCGTCTGAAGAGAGGTTGGAA	958
Qy	861	CAAGCAACGAGGATCATTTTCTTGATTCGGTCCGATCTAAGGCCAGGATTAGATCCG	920
Db	959	GAGGATAAATGCAGCAAACTTATGTCTTCATCACTGCGGACAACTTTTGTCTCG	1018
Qy	921	CGGTTCTCGAAATTTACTTCGGGAATCTGTTATCTGACGCGTCCGAGAAATGCGGCG	980
Db	1019	CCGATACCTCAAAATTTACTTTGGAATTCATAGTGCCTTGTATGTTGGATCGACTCAT	1078
Qy	981	CGAGAGCTGGTGGAGAGAGGGGTCTTCTGGCAGCTGAGTATATCCGCGGAGATA	1040
Db	1079	GAGCAACTTGTAGGAATGAAGGGTGTGCGTAGCTGCAACCGCCATCGGAGATGCTATC	1138
Qy	1041	AAAAAAGGATCAACGCAAGAGAAATATTAGAAACGGTGGAGAAATGGTCCGCGGAGATT	1100
Db	1139	CATAAGAGTTTACATGACTACGAAAGAAATTCGAGAGGAGATTGGATATCGCGCCCGA	1198
Qy	1101	CGTAAAGC-----GTTCCAGAAATCATATTTTTCGGTGGCAGATCGAGCAAG	1148
Db	1199	TCAACATCTCGCGCACCAAGGTTCGAGCGCTCATTTGTGCTGGATCCGACACCAAT	1258
Qy	1149	CTAGATCTTTACGGTCAGATTTTGATGGGGAAGCGGAGCAAGCAAGAAATATTGTG	1208
Db	1259	GTGCATGATTTGATCGAGATTTGTTGGGAAAGCTTGAAAGCATGAATCTGTTTCA	1318
Qy	1209	ATTGATGGGAGAAATATGCAATGACGCTTTGTAAGACGAGGATTTTCAAGAGGAGATTG	1368
Db	1319	ACTAATCTTCGCGCAACACTAATTTTGTATCTCTCGTCCAGAGATTTAAAGGAGCAT	1378
Qy	1269	GAGTTTGTCTGTTTGGCTAAGGACAAATGGATGCTTTTCTGCTGCTATTTTTCACG	1328
Db	1379	GAGCTTGGCATTTCTTTGGCTAAGAAATAGGATGGAGCGCATTTGCCACCATTTTACGAAT	1438
Qy	1329	GGAATTAATGGTTAATAAATGTATGTAATTAACATAATATTATTTATTAACAATTAATTA	1388
Db	1439	TTCATCAATAGTCTCCATGTGAGGAGCCCTTTGTAAGAAAAAAGTGTATCAATGTATA	1498
Qy	1389	AGTGTTCAGTAACGTGAAGA 1408	
Db	1499	AAAAGACAGACAAGTTATGA 1518	
RESULT	4		
LOCUS	AB026494		
DEFINITION	Gentiana triflora GAT106 mRNA for acyltransferase homolog, complete cds.	PLN	04-APR-2000
ACCESSION	AB026494		
VERSION	AB026494.1		
KEYWORDS	acyltransferase homolog.		
SOURCE	Gentiana triflora cDNA		
ORGANISM	Gentiana triflora		
REFERENCE	1 (sites)		

AUTHORS	Yonekura-Sakakibara,K., Tanaka,Y., Fukuchi-Mizutani,M., Fujiwara,H., Fukui,Y., Toshihiko,A., Yamaguchi,M. and Kusumi,T.
TITLE	Molecular cloning and biochemical characterization of hydroxycinnamoyl-CoA:anthocyanin 3-O-glucoside-6-O-hydroxycinnamoyltransferase from Perilla frutescens and diverse plant acyltransferase homologs Unpublished (1999)
JOURNAL	2 (bases 1 to 1622)
REFERENCE	2 (bases 1 to 1622)
AUTHORS	Tanaka,Y. and Yonekura-Sakakibara,K.
TITLE	Direct Submission
JOURNAL	Submitted (21-APR-1999) to the DDBJ/EMBL/GenBank databases. Yoshiyazu Tanaka, Suntory Ltd., Institute for Fundamental Research; Wakayama-dai 1-1-1, Shimamoto, Osaka 618-8503, Japan (E-mail:Yoshikazu.Tanaka@suntory.co.jp, Tel:81-75-962-8807, Fax:81-75-962-3262)
FEATURES	Location/Qualifiers
source	1..1622
gene	/organism="Gentiana triflora"
	/db_xref="taxon:55190"
	35..1474
CDS	/gene="GAT106"
	35..1474
	/gene="GAT106"
	/codon_start=1
	/product="acyltransferase homolog"
	/protein_id="BAA93452.1"
	/db_xref="GI:7415597"
	/translation="MAGNSEDIKVLEKRVAPPDAVEFTVPLSFFDMRWLISDAEH HUHFYRHPSPNSKFIISSIKSYSLVHLKFLPLAGNLIMPVSDSDRPELRYKRGD SVSLTAIESDMDFDLAGHQDYSYKENDLIPQLEPIVTSGDEVLPALQVTFPSN TGICIGRNHLQVGLDASSFLHFNKLWLDKNSGSLKFLPLSSLPYMSRVVQDPFH IRKLYNERKLLKSGTPTVLNPAISKDEVRATFILHPIDIMKLKFISSKRNLTGHSYNNLSFTPTVSALITWCLSKSLDTVVREKVEEDKHAANLCAFINCRQFAPIQNY FNCVLPVCMVSGSTHEQLVGNELSVNAATRAIGRAIHKLHDYEGILRGDWISPRSTA APRSTLLIYVGSARQNRHDFDADFQWGLEKHEKSVSTNPSTATLILISRRRFGALEL GISLPKRMDFATFTFNFINSLSRSL"
BASE COUNT	478 a 345 c 322 g 477 t
ORIGIN	
Query Match	14.8%; Score 218.4; DB 12; Length 1622;
Best Local Similarity	50.8%; Pred. No. 9.5e-43;
Matches	742; Conservative 0; Mismatches 661; Indels 57; Gaps 7;
Qy	3 GTGATCGAAACGTTAGAGTTGGCGCGCGCGGACTCGGTGGCGGAGCAATCGTGCGG 62
Db	62 GTTCTTGAGAAATGCGCGTGTGGCGCACACCGGACCGCGTCCGCGGATTACAGTCCCA 121
Qy	63 CTCACATCTTCCGACATGACGTGGCTGCAATTTTCATCCCATGCTTCAGCTCCTCTTCTAC 122
Db	122 CTGTCGTTTTCCGACATGCGATGTTGATCTCTGATCGACAGAACCATCTGCATTTCTAC 181
Qy	123 GAATTCCTCTTGTCCAAAGCAACATTTTCAGAAATCCATGTTCCAAACACTCAACAATCT 182
Db	182 AGATTCGCGCATCTTGTGCCAAC--TCTAAATTTATCATTTTCATCAATAATCGTCC 238
Qy	183 CTCTCTAAAACTCTCATACACTTCTTCCCTCTCTCATGCAATTTTAATCTACCCCTCA--- 239
Db	239 CTTTCCCTTGTCTCAAAACACTTTCTCCGTTAGCGGGAAATTTGATTTGCGGGTAGAT 298
Qy	240 TCCCGGAGAAAATGCGGAGTTTCGGTATCTATCCGGGACTCGGTTCTTTCCACCATC 299
Db	299 TCTCCGATAGATGCGGAGTTGCGTTACAGAAAGGAGACTCCGTTCTTTTAAACAAT 358
Qy	300 GCAGAACTAGCGACGACTTTCGATGATCTCGTCGGAATCGTCAGAAATCTCCCGTTAGG 359
Db	359 GCAGAACTAGCGATGATTTTGATATCTCGCGCGAGATCATCAGAGGGATCTTTATAA 418
Qy	360 CTCTACAATCTTCTCCCTAAATTTGCGCGCCCATTTGTCGAAGAATCCGATAGAAAATCTTC 419
Db	419 TTCAAGGATTTGATTCGCGAGCTGCCAGAACCGGATTTGTAACCTCCGCGCAGGAATTA 478
Qy	420 CAAGTTTTCGCGCGTGCAGGTGACTCTTTTCCAGCGCGGAGCGTTCGTTATGGAAATGCA 479

Db 479 CCACCTTTTGTCTTACAGGTGACGGTGTCTCCACACCGGTATATGCATTGGACGCAAT 538
Qy 480 ACGCATCACACGGTTAGCGACGCCCGTTCGTTCTCGGTTTATACCGCTTGGTCTTCA 539
Db 539 CWTTCATCAAGTCTTGGTGATGCGCAGTCTTCTTTTCGCAATTTTAAATAAATATGGGTTTG 598
Qy 540 ATGAGCAACACATTTGAAATGAAGATGAAGATGAAGATTTAAATCTTTGGCAGCTTTTC 599
Db 599 GTTGACAAATCCAATGGAGATTCATTAAGTTCCTTCCACTTCTCTCTACCTATGTAC 658
Qy 600 GATAGATCGGTCAATAAATATCCGACGAAATTT-----GACTCCATTTATTTGGGAAC 653
Db 659 GACAGATCTGGTGCAAGATCCATTTTCATATTCGTGAAATAATCTACAATGAAGAAAA 718
Qy 654 GCGTAAATTTCTTTGCCATCTGCTC-----ATCCCTCATTAACGACGACCGCAAT 707
Db 719 CTGCTCAAAATCTCAGGGCACACCTACTGTTCTTAATCAGCAATTTCTTAAGATGAAGTT 778
Qy 708 CGAACACGCTGCTTTTCCACCAATCCAAATTAAGAAATGAAGGTTGGATTCAGTCC 767
Db 779 CGAGCCACTTCATCTACACCTATTGATCATGAAGCTCAAGAAATTCATTTCGTCA 838
Qy 768 AGAGTTCCAAGTTTAGTC-----CATCTCTCATCTTTTGTAGCGGATTCGA 812
Db 839 AAAAAATCGCAACTTAACCGGTAGTAGTAATTAATATCTGCTCAACTTTCACGGTGACATCT 898
Qy 813 GCTTATATGTGGCTGGCATTAACGNAATCATTC-----ACAGCAGATGAAC 860
Db 899 GACTGATCTGGACATGCTTTGCGAAATCATTTAGACACCCGTCGTAAGAGAGAGGTGGA 958
Qy 861 CAAGACACGAGGATGCAATTTTCTTGATTCGGTTCGATTAAGGCCACGATTAGATCCG 920
Db 959 GAGGATAACATGCAGCAAACTTATGCTCTTTCATCACTGCCGACACGTTTGTCTCCG 1018
Qy 921 CCGGTTCTCGAAATTAATCTCGGGAACGCTTATCGTACGCGCTGCCGGAATCGCGCG 980
Db 1019 CCGATACCTCAAAATTAATCTTGGAAATGTCATAGTGCCTTGTATGGTGGGATCGACTCAT 1078
Qy 981 CGAGAGCTGGTGGGAGAAAGGGTGTCTTGGCAGCTGAGGTAAATCGCGCGGAGATA 1040
Db 1079 GAGCAACTTGTAGGAAATGAAGGTTGTGCGTAGCTGCAACCGCATCGGAGATGCTATC 1138
Qy 1041 AAAAAAGGATCAACGACAGAGAAATATTAGAAACGGTGGAGAAATGTCGCGCGAGATT 1100
Db 1139 CATAAGAGGTTACATGACTACGAAGGATTCAGAGAGAGATTGGATATCGCGCCCGCA 1198
Qy 1101 CGTAAAGC-----GTTGAGAAATCATATTTTTCGGTGGCAGGATCGAGCAAG 1148
Db 1199 TCAACATCTGCGGCACCAAGGTCACGCTCATTTATGTGCTTGGATCGCACAACGCAAT 1258
Qy 1149 CTAGATCTTTACGGTGCAGATTTTGGATGGGGGAGGCGGAGAAACAGAAATATTGTCG 1208
Db 1259 GTGATGATTTTGTGAGATTTTGGTTGGGAAAGCTTGAAGAGCATGAATCTGTTTCA 1318
Qy 1209 ATTGATGGGAGAAATATGCAATGACGCTTTGTAAAGCCAGGGATTTTGAAGGAGGATTG 1268
Db 1319 ACTAATCTTCGCAACACTAATTTTGATCTCTCGGTCGAGAAATTTAAGGAGCACTT 1378
Qy 1269 GAGGTTGCTGCTTTGCTTAAAGACAAAATGGATGCTTTTGTGCTTATTTTTCAGT 1328
Db 1379 GAGCTTGGCAATTTCTTTCGCTTAAGAATAGGATGGACGCAATTTGCCACCATTTTACGAAT 1438
Qy 1329 GGAATTAATGCTTAATAATCTATGTAATTAATACTAATATTATGTAACAAATTAATA 1398
Db 1439 TTCATCAATAGTCTCCATGTCGAGGAGCCCTTTGTAAGAAAAAAGTGGTATCAATGTATA 1498
Qy 1389 AGTGTGAGTAACGTGAAGA 1408
Db 1499 AAAAGACAGACAAGTTATGA 1518

RESULT 5

E12757 1508 bp DNA PAT 24-JUN-1998
LOCUS Senecio cruentus mRNA for acyltransferase,,partial cds.
DEFINITION E12757
ACCESSION E12757
VERSION E12757.1 GI:3251589
KEYWORDS JP 1997070290-A/5.
SOURCE unidentifed.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1508)
AUTHORS Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y.,
Yonekura,K., Mizutani,M. and Kusumi,T.
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 5 18-MAR-1997;
SUNTORY LTD
COMMENT OS Senecio cruentus
PN JP 1997070290-A/5
PD 18-MAR-1997
PR 30-JAN-1996 JP 1996046534
PF 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO,
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAHI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
C12R1:865),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1508
FT /organism='Senecio cruentus'
FT /tissue_type='petal'
FT /clone='pCAT48'
FT 1..1367
FT CDS /product='acyltransferase'.
FEATURES
source
1..1508
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 442 a 293 c 296 g 477 t
ORIGIN
Query Match 12.9%; Score 190.6; DB 10; Length 1508;
Best Local Similarity 49.6%; Pred. No. 5.7e-36;
Matches 676; Conservative 0; Mismatches 654; Indels 33; Gaps 6;
Qy 2 CGTGATCGAAACGTTAGAGTTGGCGCCCGCGGACATCGGTGGCGAGCAATCGGTGCC 61
Db 5 CATCTCGAACATGCCGAATATCGGCCCTCGGGCACCATCGGCCATCGCTCGTTATC 64
Qy 62 GCTCACATTTTCGACATGAGTGGTGCATTTTCATCCCATGCTTCAGCTCCTCTTCTA 121
Db 65 TCTTACTTTCTTCGACATTAATTTGCTACTCTCTCCCTCGGTCCACCATCTTTCTTCTA 124
Qy 122 CGAATTCCTTTGTCGAAGCAACATTTTTCAGATCCATCGTCCAAACACTCAAAACATC 181
Db 125 TGACTTCCACATTTAAATCCCATTTTCATGGACACTATTGTTCCAGGCTAAAACAATC 184
Qy 182 TCTCTTAAACACTCTCATACACATCTTCTCCCTCTCTCATGCAATTTAAATCTACCCCTTCATC 241
Db 185 TTTATCGTCACTCTTCAACATTTTTCCTCGTTTTCCTAGTAAATTTGATTTGTTTCTTAA 244
Qy 242 CCCGG-----AGAAATCCCGAGTTTCGGTATCTATCCGGGAGTCCGGT 286
Db 245 CACTGATGGTTTCGGGTTTAAATAAAAAACAGAAATAAAACACGTTGAAGGTGATTCTGT 304
Qy 287 TTTCTTACCATCGCAGAACTAGCAGACTTCGATCATCTCGTCGGAATCGTCCAGA 346
Db 305 TGTGGTTACTTTTCGAGAAATGTTGCTTGAATTTTAAATAATTTGACAGGAAATCATCTCG 364
Qy 347 ATCTCCCGTTAGGCTCTACAACTTTTGTCCCTTAAATTTGCCGCCCAATTTGTCGAAGAATCCGA 406

Db	365	AAATGTGAAACTTTTATCCACTTGTACCTTCATTGGGAAATGCAATCAAAATTTATGTGA	424
Qy	407	TAGAAAACTCTTCCAAAGTTTTCGCCCTGCAGGTGACTCTTTTCCAGGCGGAGCGCTCGG	466
Db	425	TTCCGTCGCGCTCCCACTTTTTCACCTTCAAGTGAGCTTTTTTCCGGCTCGGGTATATC	484
Qy	467	TATTTGAATAGCAACGCATCACACCGTTTAGCGACGCCCGCTCTTCTCCGCTTTATAAC	526
Db	485	ACTAGGAATGACGAATCATATAGCCTTGGTGACGCTAGCACGCGTTCAACCTTTTGAA	544
Qy	527	GGCTTGGCTTCAATGAGCAACACATTTGAAATTAAGATGAAGATGAAGAAATTTAAATC	586
Db	545	AGGTGGACTTCGATTATTCATCTCGTGTAGATCGTCTTTTTTAACGAAAGG---ATC	601
Qy	587	TTTGGCAGTTTTCGATGAGATCCGTCTATAAATATCCGACGAATTTGACTCCATTTATG	646
Db	602	TCACCGGTTTTTGTATGATGAT---TGATTAACATCCACATTTTAGATGAATAAGTTGAG	658
Qy	647	GAGAAACGGGCTAAATTTCTTTGCAATCTCGTCATCCCTCATACCGACGCGCGCAT	706
Db	659	ACATACAGGCTCGAAGTTTTTATAACCTTCGAGCCTTGTGGTCCCACTGATAAGT	718
Qy	707	TCGAACACAGTTCGTTTTCACCAATCCAAATTAAGAAATTAAGAGGTTTGATTCAGTC	766
Db	719	TCGGTCAACGTTTGTGTGACCGCGAATTAATATCAATCTACTAAAGAAAGGCTCTTAAC	778
Qy	767	CAGAGTCCAAAGTTTAGTCCACTCTCATCTTTTGTAGCGATTTCAGCTTATATGTGGC	826
Db	779	CCAAGTGCCNAACCTTGGATGATGTCATCTTTTACGGTAACCTTGTGGTTTATATGAG	838
Qy	827	TGCATAACGAAATCATTC-----ACAGCAGATGAAGCAACAGCAACGAGGATGCATT	880
Db	839	TTGCATAGCGAATCACTCGTAAATATAGGAGAAAGAAAGGCGAGACGAGTTAGAAC	898
Qy	881	TTTCTTGATTCGGTTCGATCTAAGGCCACGATTAGATCCGCCGGTTCCTGAAAATTA	940
Db	899	GTTTCATAATCACCATTGATGTGCGATCTCGTCTTGATCCACCAATTCACACGCTACT	958
Qy	941	CGGAACCTGTTATCTACCGCTCGCGAGAAATCGCGCGCGAGACGTGTTGGGAGAA	1000
Db	959	TGTAACCTGTGGTGACCACTGTCTCCGACCTTAAAAAATGTCGTTTGTAGTACGCAAA	1018
Qy	1001	AGGGGTGTTTCTGGCAGCTGAGGTAATCGCGCGGAGATAAAAAAGAGTCAACGACAA	1060
Db	1019	TGGGTATGCACCTTGGTCTAAAGTAATTTGGAGAGTCTATATGCAAAATGATATATA	1078
Qy	1061	GAGATAATTAGAAACGGTGAGAAATGGTCCCGCGAGATTTCGTAAGCGTTGCAGAAATC	1120
Db	1079	GGACGGAATCTTGAAGATGCCCGGAGATGGCATGA---ACCTTTTCATGATCCCGGCTAG	1135
Qy	1121	ATATTTTTCGGTGGCAGGATCGAGCAAGCTAGATCTTTACGGTGCAGATTTTCGATGGG	1180
Db	1136	GAAGATTTGGTGTGCTGGTACACCTAAGCTCAACTTGTACGACTTTGATTTGGGTGGG	1195
Qy	1181	GAAGCGGAGAAACGAGAAATATTGCGATTGTATGGGGAGAAATATGCAATGACGCTTG	1240
Db	1196	GAAGCGGATAAAGTATGAGACTGTTTCAATAGA---CTAATAACGTCGATTTCTATAAA	1252
Qy	1241	TAAAGCCAGGGATTTCGAAGGAGGATTGGAGGTTGCTTCTCTTTGCCCTAAGGACAAA	1300
Db	1253	TGCAAGCAAAACATCAGCACAAGATCTTGAAATTTGGATTGAGTCTACCGAGTATGCAAT	1312
Qy	1301	GGATGCTTTTGGCTTATTTTTCACCTGGGAATTAATGGTTAA	1343
Db	1313	GGAGGCGTTTTCAGCATCTTTGATGAAGGATTAGAGTGCAA	1355
RESULT	6		
LOCUS	AB010708	1679 bp	mRNA
DEFINITION	Gentiana triflora mRNA for Anthocyanin 5-aromatic acyltransferase,	PLN	20-FEB-1999
	complete cds.		

ACCESSION	AB010708	VERSION	GI:4185598
KEYWORDS	Anthocyanin 5-aromatic acyltransferase.	SOURCE	Gentiana triflora petal cDNA to mRNA, clone:pgAT4.
ORGANISM	Gentiana triflora		
REFERENCE	1 (bases 1 to 1679)		Eukaryota; Virdiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.
AUTHORS	Fujiwara,H., Tanaka,Y. and Kusumi,T.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-JAN-1998) to the DDBJ/EMBL/GenBank databases.		
MEDLINE	Hiroiyuki Fujiwara, Suntary LTD., Inst. of Fundamental Research;		
FEATURES	1-1-1, Wakayamadai, Shamamoto-cho, Osaka 618-0024, Japan		
source	(E-mail:Hiroiyuki.Fujiwara@suntary.co.jp, Tel:+81-75-962-8807, Fax:+81-75-962-8262)		
CDS	2 (sites)		
	Fujiwara,H., Tanaka,Y., Yonekura-Sakakibara,K.,		
	Fukuchi-Mizutani,M., Nakao,M., Fukui,Y., Yamaguchi,M., Ashikari,T.		
	and Kusumi,T.		
	cDNA cloning, gene expression and subcellular localization of		
	anthocyanin 5-aromatic acyltransferase from Gentiana triflora		
	Plant J. 16 (4), 421-431 (1998)		
	99097837		
	Location/Qualifiers		
	1..1679		
	/organism="Gentiana triflora"		
	/db_xref="taxon:55190"		
	/clone="pgAT4"		
	/tissue_type="petal"		
	6..1415		
	/codon_start=1		
	/product="Anthocyanin 5-aromatic acyltransferase"		
	/protein_id="BAA74428.1"		
	/db_xref="GI:4185598"		
	/translation="MEQIQMVKLEKQVTPSDTVDVLSLPVTFDDIPWLHLNMQ		
	SLFYDFPYRTHFLDTVIPNLKASLSLTKHYVPLSGNLMLPIKSGENPKFOYSRDE		
	GOSITLI VAESQDFDY LKGHQLVDSNDLHGLFYVMPRVIRTMQDY KV IPLVAQVTV		
	FNRG IAVALT AHSIADAK SFVMP INAWAY INKFGKDADLLSANKLPDFDRS I IKDL		
	YGLETFNEMQDVL EMFSGKPPREKNVATYVLSLAE IOKLKNKVLNLRGSPPT		
	IRVITFTTCGIVTWCWKRSKDDVVSESSNDENELEYEFSFTADCKRLTLTPCPNPF		
	GNLASCVAKATHKELVDKGLLVAVAAGIEATIEKRLHNEKGLVADAKTWSLSENGIP		
	KRFLGITGPKFDSYGVDFGKPAKFDITSVDYAEIIVYIQSRDFEKGVEITGVS		
	SLP		
	polyA_site	1679	
	BASE COUNT	488 a 350 c 359 g 482 t	
	ORIGIN	/note="34 a nucleotides"	
	Query Match	11.6%;	Score 172.2; DB 12; Length 1679;
	Best Local Similarity	49.5%;	Pred. No. 1.8e-31;
	Matches	701; Conservative	0; Mismatches 668; Indels 48; Gaps 8;
Qy	48	GAGCAATCGCTGCGCTCACATCTTTCGACATGACGCTGCGCTTCATCCCATCCTT	107
Db	78	GAGTATCGCTACCGTGAACATCTTCGATATCCCTGGTTCGACTTGAATAGATCGAG	137
Qy	108	CAGCTCCTCTTTCAGAAATCCCTTGTTCACAGCAACATTTTTCAGAAATCCATGTTCA	167
Db	138	TCCTTCTGTTTACGACTTTCGCTACCCCAAGAACACATTTCTTGTATCCCT	197
Qy	168	AAACTCAAAACATCTCTCTTAAACCTCTCATACATCTTCCCTCTCTCATCAATTTA	227
Db	198	ATCTTAAAGCCCTCTTGTCTCTCATCTCTAAACACATACCTTCGCTTACGGAAATTC	257
Qy	228	ATCTACCCCTTCATCCCGGAGAAATGCCGAGTTTTCGGTA-----TCTATCCGGGAC	281
Db	258	TTGATGCCGATCAATCGGGGAAATGCCGAGTTTTCAGTACTCCCGCTGATGAGGGGAC	317
Qy	282	TCGGTTTCTTTCACCATCGGAGAAATTCAGGACGACTTCGATGATCTCGTCGGAAATCG	341
Db	318	TCGATAACTTTGATCGTTCGCGACTCTCGACGAGATTTTGACTACCTTAAAGGTCATCAA	377


```
SNLRLYFYPRYGSQNAVKSFFSVKVNQGTLLNNSADLTWKASKPQTEFIKEFI
IPVQTLNLFTSLDSLAFFNGIEIVSIENREYSKGGFDDVITNNGSSVDFHIENST
AFETVYRLNGTSGVSGMFRWRWSDDEIILSESSGISIPVDIKINYTEKTPSYVA
PDDVYATSRMGADHPQONLNLWLTFTVDAGFSYVLRHFCETLSEVKNKEGQRFV
SIFTEHOTATDWDVPMGSGSMIPMYLDYTVIAGSGSGRHDLRLDPLVSNPKY
YDALLNGVEILKNDPDCNLAGNPDLVSPDLIPNRPATRIKKNSHIIPITLAVVG
SLVAMFVGVGLVIMKKKSKSPSTNSWCPLPHGTDSTNTPAKSLPADLCRRFSI
FEKTSATNDEGLIIGVGGFSYKQIDGGATLVAVKREITSNQAKFEFTELEM
LSKLRLHVLWSLIGCYCEDENMLVLEYEMPHGTLKHLFRDKTSOPLSWKRRLEIC
IAGARGQLVHTGAKYTI IHRDITKTNILDENFVTKVSDFGLSRVGPTASOTHVST
VVKTCFGLDPEYRQVLTAKSDVYSGFVVLVLELCCRPIMQSVPEPEQADLIRWK
SVNRGIVDQILISDSADITSLEKFCIEIARCVDQDERPMPMDVNVVLALEFALQ
LHETAKKNDNVSZSLDLMPSGEVGTITDDGDDLSRTTGHVGKSTTTDDSDVVGDER
SGSGWGVFSINEPKAR"
6701..7210
/unknown="gene_id:MXF12.3
/known protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BABI0825.1"
/db_xref="GI:10177546"
/translacion="MINLEKEAKEKOILLIRGKRKIVHHVNTLLVKIYEASAIIDORH
LFLKPYGPNYFIKAWLDDTKYGYVVKDSFKVNPQPLKIPLNPAPQILVLELLK
GMSTRDPGTNRVTVMGRAKIRUSPLMTSYKERPISGGIDLVLGNSDQCVVKKGYLY
AMKLHRSAD"
7670..10111
/unknown="gene_id:MXF12.4"
/codon_start=1
/evidence-not_experimental
/product="receptor protein kinase-like protein"
/protein_id="BABI0826.1"
/db_xref="GI:10177547"
/translacion="MNCNVLFLLSVLVSVTAGVTAAYHPTDVFLENGGDTSNVNDNSG
RNTWTSRQIILSNLNVASTSEASQKAGVSRIPYMKARIEFSEFTYSPVTPGSLF
LRLYFYTQKSGFDVANSFYSKYNGFTLLRNFNADSTVQASIPLSNLIKELIPV
HQTLNLFTFSKLLAFNGIEIVSPDRFYSKGGDNLRNVNSDVFDFDINDSTAFE
SVHRLNVGGQIVNEVDSDGFRWLSDSFGNSGSIIVPGVKINTKPTKYNSTVAYD
VIATSRLMGNSLNLFTMGFTLDAGYNLVLRFHFCETLPQVTRAGQRFVSEFED
KMAKETDVLIRSGGPRIPIYLDVSVYGFESGMIOPELRLDLVPLKDTNQTVDAIL
SGVEILKNDSDGNLARNPELLAVSTDSTPDDSNVTPPIKKGPHLVILILVSGSVIG
LATFIVIMLLIRQMKRKNKENSVMFKLLKQYIAELKKITKTSFSTHTVKGSGFG
TVTRGNSUNGTRVAVKVLKNGDDFINVETMSQTSNHNVSLLGLFCYEGSKRAI
ISFELHGSDDQFISRNKSTPNVTLDYIALGARGLEYLHYGCKTRIYVHFDPKQN
ILDDNFPKRVADFLAKCEKRESILSLDITRTGTYIAPEVVSRYMGISHKSDYV
SGMLVLDIMIGARKVETTCNGSTAYFPDWIYKOLENGDQWTWIGDEINEEDNKIVK
KMLVLSLWCIRPCPSDRPPNKKVMEIEGSLDALELPKPSRHSIHSTELVLESSLSDG
QEAERQITLDSII"
10819..13239
/unknown="gene_id:MXF12.5"
/codon_start=1
/evidence-not_experimental
/product="receptor protein kinase-like protein"
/protein_id="BABI0827.1"
/db_xref="GI:10177548"
/translacion="MICFILFVFSFLVVSATAPYKPDVDFLINCGETDVPFDNHGRT
WTOEBKILPKNSDNASFSVWSYKESGIPQVPMYARTAFRSDTYSPVSPGKWFLL
RLVYFPTSYSKSGDAVNSFVSVNDFTLQNEASDLTVKASIPESKSLIKELIPVY
LTNLNLPSPNSNLAFFNGIEIVSPDRFYSKGGFDDLTNNGSLDFIDNSTASET
VHRLNVGGHVDENSGMFRWLSDSDFELIGGVSPDPDMNISYTEKTPAYVAPAY
VYTCRMMQNAQDTYLNLFNLWLTFTVDAGFSYVLRHFEKYLANKANORVFSIFLG
NOMAREMDYIRLSGPRIPYILDFTRIYVSGSGPRDLRLDHLPLVKONPEYFEAL
NGVEILKLNNSGLAIQDNEKLPNPLSSNLTPNVTQOIKGSSHLVKIFITAVGP
GTGLATPVVYVLMWLMQMRKRNKEERVNFKLLNMYTVAELKKITKTSFSYIIGKGG
FCVYVGSNLSNGKVAVKLDLKGSAEDFINVASMSTSHNVISLLGFCFEGSKR
AIYVEFLNSLQOFSRNRKSLQDVTLLIAGLARGLEYLHYGCKTRIYVHFDPKQN
QNILLDNLCPKVSDFGLAKCEKRESVLSMDTRGTIGYIAPEVSRMYGRVSHKSD
VYSFGMLVIDMIGARKEIVETPDWISASSTYFPDWIYKOLEGDEQWTFIDEITKEEK
EIAKMLVIGLWCITQPCPSDRPPNKKVMEIEGSLDALEIIPKPSRHSIHSTEVITESS
LSDGGEV"
join(16412..16498,16859..16924,17003..17038,17085..17192,
17337..17461,17541..17640,17803..17974,18073..18191,
18285..18395,18488..18632,18712..18844,18934..19009,
19232..19297,19385..19492,19591..19698,19873..20070,
```

```
20164..20373)
/unknown="gene_id:MXF12.6"
/codon_start=1
/evidence-not_experimental
/product="ABC transporter-like protein"
/protein_id="BABI0828.1"
/db_xref="GI:10177549"
/translacion="MGNKKLITGSSKTHGSSSYRDLQNOEDKPKANGSENGLD
LEHGVAAVNVGFRVFALICQYISIMVSOAKPDAGLVGTIGTALLIGTSTNLLVPKF
GGMILDIVSDRVKTPQOETSLIAVRNAVVIILLVVGISICTALRAWLFNSASERVV
VLMIDIVSDRVKTPQOETSLIAVRNAVVIILLVVGISICTALRAWLFNSASERVV
ARLRKDLFRHLMHQETAFDYDTYTKGELLRLSLEDTOIKNAAATNLSALRNVTALI
GVMFTYSWKLTLALLVVPVIVSAVKQFGRYLRSLSHSTTQAAAAVAASIAESFGA
VTRVFAKESYMVSYQSKVDETLLKGLKQAVLVGLFGGLNAAFTLSVTVVSYGA
YLTICGSMVTGALTSEFLYSYVSSVSSLSYTTAMKAAGASRRVFOQLDRVSSMS
SGDGKVPNPDGDELNDVWFAYPSRPSHMLKGLISRLTSGKSVKALVPGSGGKKT
IANLIRFYDPLAGKILLNGVSLMEITSHOYLHKQISVSOEPIIFNCSVEENIAYGPD
GEASFTDIENAAKMANAHEIFAPPKYNTVVGRLRLSGGQKQRIARALLTNPS
VLILDEATSADESEYLVQDAMDLSMAGRTVILVIAHRLSTVTKTADCVAVISGGEVAE
KGTHELLSLNGIYTNLKRQLOSSSSVTTL"
join(25348..26359,26401..26756)
/unknown="gene_id:MXF12.7"
/codon_start=1
/evidence-not_experimental
/product="anthocyanin acyltransferase-like protein"
/protein_id="BABI0829.1"
/db_xref="GI:10177550"
/translacion="MVNEEMESSLKVIDVARVTPSNSDSESLLTLPITFDLLWYKLLH
AVERVIFYKLTASRPFFDSVIVPNLKTSLSSLSHYLPLAGKLVWEPLOPKIVYT
PNDVSTFAESNADFSRLTGKEPFTTELYPLVPELHVSDDSASVFSQVTLFPNGG
FCISVNAHVAVLGKTTTNFLKSWARTKNQDSFLPDQLIPVSDRTVTKDPMDLDTKI
LNAMHRVAKYFTGKKEPNSKLKLSMSPGDPVFRYTLNLTREDIQRLERLKKES
SSSSVSSPKELRSTFVIYVSYALTCLIKARGDPSRPVGYGFVDCRSLMWPPVPS
SYETPMSEGEFLAAARMVSDSVEALDENVALKIPEILEGFTTLLSPCTQVLSVAGSTR
FGYGTGDFGWRGKVVVSDIDGAEISFAESRDGSGVGLFSLKHEMDVLVDLLH
KGLN"
complement(join(27466..27879,28046..29073))
/unknown="gene_id:MXF12.8"
/codon_start=1
/pseudo
/evidence-not_experimental
/product="transposase-like protein"
complement(join(30804..30874,30947..31289))
/unknown="contains similarity to Ac-like transposase"
gene_id:MXF12.9"
/codon_start=1
/pseudo
/evidence-not_experimental
32432..33823
/unknown="gene_id:MXF12.10"
/codon_start=1
/evidence-not_experimental
/product="anthocyanin acyltransferase-like protein"
/protein_id="BABI0830.1"
/db_xref="GI:10177551"
/translacion="MVSSLNIEIIVARVTPSNSDSESLPLTYFDLIYKLAVERV
IFRYNTNPPFFDSVIVPNLKTSLSSLSHYLPLAGKLVWEPLOPKIVYSONDDV
SFSAVATNADFSISLGNPEPPELPLPALOSSDASISVFSQVTLFQNGECIGV
SAHVAULDKTTTFLKSWAICKHQDFSLQDLITPYDRTVTKSPDSENKVLNWR
SFTKILAGGCEPANPSKLNPSPFEIGDPVVVTLQTRDIEDIOTLERLKRREVSSES
STSSKRLMTSTFVIYVYLVCIIRARGEPHPRVGYAFSDVCRSLMNPPTNYFGN
CIAGSRMLMTAKMEIEGLAAATMVSDSIEWDESFAMKIPDFVAYATLPETOL
ILVSGSNRFQVYELDFGWRGKPVVSVISPGNGISMAESRDQNGSVEIGFSLKKHEM
DTLIDLHLCVETI"
34363..35709
```

Query Match 7.5%; Score 110.8; DB 12; Length 66237;
Best Local Similarity 46.5%; Pred. No. 3.1e-16;
Matches 616; Conservative 0; Mismatches 662; Indels 48; Gaps 6;

Qy 37 ACTCGGTGGCGGACCAATCGGTCCGCTCACATCTTCGACATGACGCTGCATTTTC 96
||||| ||||| | | | ||||| ||||| ||||| ||||| | ||||| |
Db 34418 ACTCGTGGCGGTCACTCACTCTCCGCTCACTTTCTTTTGACCTCTCTGCTCAACACA 34477

JOURNAL	Submitted (15-JUL-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE	7 (bases 1 to 101176)
AUTHORS	Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S., Kim C., Altafi H., Bel B., Chin C., Chlou J., Choi E., Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M., Palm C., Pam P., Sakano H., Schwartz J., Southwick A., Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A. and Ecker J.
TITLE	Direct Submission
JOURNAL	Submitted (19-JUL-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
COMMENT	On Jul 15, 2000 this sequence version replaced gi:2618677.
FEATURES	Location/Qualifiers
source	1..101176 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="1" /clone="F21B7"
CDS	join(736..1038,1276..1713,1874..2119,2205..3032,3117..3274,3483..3610,3725..4134,4313..4435) /note="unknown protein; similar to ESTs dbj AV529515.1, dbj AV529266.1, dbj AV554950.1, dbj AV523761.1, dbj AV523761.1, dbj AI996703.1, and dbj AV546743.1" /codon_start=1 /evidence=not_experimental /product="F21B7.1" /protein_id="AAF86508.1" /db_xref="GI:9280639"
CDS	/translation="MQPLPARSGDEGFWMNSHPLLVAAGDETNGTGLGHFSONGSLARGSDSKAGDAINPTTVFYSLRSHSVYVILFRFSSVCMIRCSRVAVVGLANGIYCDVALLENKFSVLTVPQVRQGTIRVNGYGMVAGFRLVYASKSMWTKTORLPDFTSPSPSLSPSSSGSSGMARYAMESKOLANGLNLDGMGYKTKLSKYCDMLPDGSTPASPNAIKVGVSGSDAENAGVAVKDLVSGALVSQFKAHTSPISALCFDPSGTLTVAVCGNNINVOIIMPSSHNPAGDLSPEWESSHVHLFKLRHGTTSATVODI CF50SQWVAIISKSGCHIFVLNSSGSDAFOCEGEETRLPASLPMWFTQSLSNQSLSPTVALSVSRIKYSFSGWLTVSNATATGKVFVPSGAAVAFKSVTHDLQNSTNALEHLIYTPSGHVVOHELLFSVCETSPENGLRVOKTSHVQVQDDLRV KVEPIQMDKSDWLEETELRPKSIETEKQYDLETVSNHLTSHEDACLDMNSHFS EDKYKSCSKPERSHCYLSNFVKYTGMLPWQNSKISFHVMDSPRDSSTGTGEF EIEKYPAAHEIIEIKQLLPVDFHFKATLEDKRFMSKYHTSATGSHQVNGKICQDI INCHSKPGTSEASEEGSTKOMENLHSDHMSNISKSLPLYPTVNGIYKETEKN ANGMKXPYAKLSLTKETRTNGFTTPIILTQSVNQMLSTGCKPPMGFGFALHEHC KAVADPKEHLKKLDEVTNVHLNVNNTFELQGDKNVNSQVLAFAFGCILYKDFL RVGNNEFVFPLLDYKIWLKK"
CDS	complement(4838..6223) /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F21B7.2" /protein_id="AAF86541.1" /db_xref="GI:9280672"
CDS	/translation="MASCIELHFTHLHIPVTINQOFLVHPSSPTPANQSPHLSLYLS NLDDIGARVFTSVSYPTNNRESVLKRLQDALSEVLVPPYPLSGRLREVENGLK EVFGEQGLVMTVANSMDLADGLTVNPAMLPLIFRNGPEAYKILEMPLLIAQ VTPTCCGFSGLJRLCHICDGFCAQMOFLGWAATKGLIADPEVPWDRFTKPN PPWVKYPHHIEVLPERSLNTSLMDTKLQKCYRISKFCQCRVKSSTAQGEDPTLVCSTFDMAAHIRWSWALDKVPLDPLNDLRTFSVNVTRLETLLKLRKSGVNVVCLACAM SVSEVINDLSLSTRLVQDARLRVSDYVDVVRKPRLEFGKLTITQWTR FEMVETADFCWGKPVYAGPIDLRPTPOCVLLPQGGVSGNQDSWVCLCLPPTAVHT FTRLLSLDNHK"
CDS	complement(join(6700..7357,7482..7884,8620..8839,8920..9241,9758..10380)) /note="unknown protein; similar to ESTs gb T04168.1, gb T21964.1, and gb AW004462.1" /codon_start=1 /evidence=not_experimental /product="F21B7.3" /protein_id="AAF86540.1" /db_xref="GI:9280671"
CDS	/translation="MGHDSFCYLIVLCALRCGIIALMQICALQKERRSKMSSDRS SQADEFETGKVLGASGKEIPAMFHTPPDPTLTSLKOTAPPQQTLTPTVDLKGK SMDLISRSVVEKIGDAAEWGFQVNVHGISVEVMERKGEIRRFHQEDQDEVKRFY SRDTRDVLVYSNIDLHTCNKAANRDTLACYMAPDPKQLQDLPAVCGETIMWEYSKOL MTGEFPELLSEALGNPNHLKDMGCAKSHIMFGQYPPCPQPDILGLTSKHTDFS IITLLQDNIGLOVHDCQWVDVSPFGALVINIGDILLQSLNDKFIISAEHRJANGS SEPRISCFVSTFEMKPNRIYGIKELLSEQNPAKYRDLTITEFSTFSQTLRIE IPALFRATPATLASLKSPPPKHLIPTVDLKGASVVEKIGEAEEKGLFHLVNHGIP VEVLERMODIGRFGHEOEPAKRFYSRDRDVLVFSNHDLQNSEAKSDMDTLCGYT APERPRLDPAVCGEITMLESKIMSILGERLFEELSEALGNLSHLKMDCAKSOYM VGQHYPCQPDILGINKHDDISFLVLLQDNVGGLOVHEQIWDVTPVPGALVIN IGDFLQVKIMLLRSCFISCMIWGLLILFYQDLITNDKFIISAEHRJANGSEBPT SVAIVFTFMRAYSRYGPIKDLLSAENPAKYRDLTITEFSTFSSKTLDPACKLHHEK I"
CDS	10833..11348 /note="unknown protein; similar to EST embj Z30724.1" /codon_start=1 /evidence=not_experimental /product="F21B7.4" /protein_id="AAF86509.1" /db_xref="GI:9280640"
CDS	/translation="MRGFSVSSWLTTSVEIGEVFLGSTWLDVAVAEADVHWVES ACLMVKLDLRLWLRVPVAGLSEPSFFSLRFLSPSIRIEGCFAPKFSDLRWRSE LQSCGVVLRWGDLQIGWRSVGNWVSVRFVAGPASSLVRDLRODYSANDLVFVVV LQSSGCPIRR"
CDS	join(11965..12033,12173..12303,12448..12519,12596..12685,12756..12915,12985..13041) /note="similar to ATHP1 dbj BAA37110.1" /codon_start=1 /evidence=not_experimental /product="F21B7.5" /protein_id="AAF86510.1" /db_xref="GI:9280641"
CDS	/translation="MWTIVVAQLORQDQIVISLYQOQFLDNQFSELKRLQDEGTPDF VAEVSLUFFDDCSKLINTSILSRPNVDFKQVDGSHQLKGSSSLCFLVLSGARR VKNVCISFECCDVQNRREGYSSNHTNCDNDIWHRCRLCQQVDYEEKMLKTKLQDLF NVSFKLCENLVLLLEKILQILQAGGTIPQVDIN"
CDS	16495..17682 /note="similar to protein kinase pir A57676; similar to ESTs gb A1993651.1, dbj AV538995.1, dbj AV522524.1, and dbj AV538995.1" /codon_start=1 /evidence=not_experimental /product="F21B7.6" /protein_id="AAF86511.1" /db_xref="GI:9280642"
CDS	/translation="MRTFKLVWCLMFLLRFGFTTEALIDPVDFLALQAIRKSLDDLPG SKFESWDFTSDCGFAGVYCNQDKVLSNLGDDPRAGSLGSRIDIPAIKLSALTET SIYPGRIMGALPATISQLKDLRLAISRNFIISGIPASLGAEVGRGLRTLDLSYNLQTLT IPSIGSLPELSNLIILCHNHITGSIPLPQLTIRDLKRNSITGSIASLSPSLQY LSLAWNOLTGVSYHVLLRLNQLNYLDLSLNRFTGTIPARVFAPITNLQLRNPFEG L IOPANQVTISTVDLSYNRESGSIPLSSVENLYLSNRFTGEVPASFEVERLLSANIQ TLXLQHNFLTGIQISPAAEIPVSSSLCQVNCMVPPLQTPCPLKAGQKTRPTTQCTE WRG"
CDS	complement(join(18462..18515,18612..18742,18840..18975,19064..19139,19214..19410,19497..19595,19668..19752,19865..20103,20179..20253,20345..20649,20798..20902,21088..21216,21356..21437,21513..21608,21780..21849,21929..22025,22408..22502,22586..22659,22810..22888,23041..23101,23206..23315,23411..23556)) /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F21B7.7" /protein_id="AAF86539.1" /db_xref="GI:9280670"
CDS	/translation="MPEKFKRRSNFIICLPKKECARGOILGYKMETKAIGFHHLSR PQIVCGHTLTAVFNHNSHQILILFGGSTAVANNHNSLPEISLDVTNVSFSDVLT R KTRLPNIPGVDPSPRACHAALYLLIIQGIQGPSGSDGVVFMLOMTNNKWKFLV GGTPSPRYGVHVPMDIAAORWLVIISGNGNEIILDDTWALDRGFSNDRLNPSCNQPS GRMYASSSREDGIFLLCGGIDHSGVTLDGTGYLKMDSNDVWTPVPAVAPSPRQHTA VFGSKLHVIGGILNRARLIDGEAVADTNQTPETSASGANRQNOYQLMRKCHHAASF

Query Match	7.5%	Score 110.8	DB 12	Length 101176
Best Local Similarity	47.2%	Pred. No. 3.3e-16		
Matches 651	Conservative	0	Mismatches 692	Indels 35
Qy	2	CGTCATCGAAACGNGTAGAGTTGGGCGCGCCCGGACTCGGTGGCGGAGCAATCGGTGCC	61	
Db	37273	CATCATCGAGACCTGTCACATCTCTCCCCCAAGGGCAGCGTGCATCCACCATCTCTCC	37332	
Qy	62	GCTCACATCTTCTTGACATGAGCTGGCTGCATTTTCATCCCATCGTTGAGCTTCCTCTCTTA	121	
Db	37333	TCTCACCTTCTTCGATGCCCTCGCTCTCTCCCACTCGCCGATTTCTCTCTCTCTT	37392	
Qy	122	CGAATTCCTTGTTCGAAGCAACATTTTTTCAGAATTCGATCGTTCCAAACATCAACAATC	181	
Db	37393	CTCTTAGCAAACTCAACTGAATCTTCTCTCAAGACTTTGTACCAACCTCAACAATTC	37452	
Qy	182	TCTCTCTAAAACCTCTCATACACTTTCTCCCTCTCTCATGCAAAATTAATCTACCCCTTCATC	241	
Db	37453	CCTCTCCATCACTCTCAACATTTCTTCCCTTAGCCCGGTAACTGATTTCCCGCCTCG	37512	
Qy	242	CCCGGAGAAAATGCCGGATTTTCGGTATCTATCCGGGGACTTCGGTTCTTTCACCATCGC	301	
Db	37513	TCCCGACCTCCATATTTACACTACAAGCTGGCGAAGACTCTCTGTTTTCCAGGTAGC	37572	
Qy	302	AGAATCTAGCGA----CGACTTCGATGATCTCGTCGGAAATCGTCAGAATCTCCGTTAG	358	
Db	37573	AGGTTCTACTGAAACCGACTTCGACCAACTGAAGTCTGATTTCCCTAAAGATATCAGTGT	37632	
Qy	359	GCTCTACNACTTTGTCCCTTAATTTGCGGCCCATTTGTCGAAGAATCCGATAGAA---A	415	
Db	37633	GTTGCATGGCGCTCTGGCCAAAGTTACCTCTCTTCCCTCAGCTCTCTCCGAAGAATTCAGAT	37692	
Qy	416	CTTCCAAAGTTTTCGCCGTGAGTGACTCTTTTCCAGGCGGAGCGTTCGGTATTGGAAT	475	
Db	37693	CGGACCAATATGCAATGCAAGTCAACCTTTCCCTGGAGCTGGAATCTGTATAGGCA	37752	
Qy	476	AGCAACGCATCACACCGTTAGCGAGCGCCCGTCTTCGCGTTTATAACGGCTTGGTC	535	
Db	37753	CTCAGCTACACATGTTGTAGCAGTGGAGTCACCTTCAGTCTCATTTTCATGAAGTATTGGAT	37812	
Qy	536	TTCAATGAGCAACACATTTGAATTAAGATGAAGATGAAGATGAAGAATTTAAATCTTTGC	591	
Db	37813	GTCAITGACCAAAATCCAGCGGCAAGATTCGGCCACGGTTCTTCTACGCTCTCTGCGCTAT	37872	
Qy	592	---CAGTTTTCGATAGATCCGTCATAAAATATCCGAGCAAAATTTGACTTCCATTTATTGGA	648	
Db	37873	TCACAGCTGCAGAAACATAATCAAGGCCAGGAGAGGTACCCGACGACATTTGGAGCG	37932	
Qy	649	GAACGGCGTAAAATTTCTTTTGGCAATCTCGTCATCCCTCAATACCGACGACCGCATTC	708	
Db	37933	GTTTTGGAGCCAAACACTGCAAAACACAGTTTCATGTTACGCTGA-GAACAATGGTCA	37991	
Qy	709	GAACACGTTCTGTTTTCACCCAATCCAAATTAAGAAATTAAGGTTGGAATTCAGTCCA	768	
Db	37992	GAGCCACTTTTACATTTGAGCCGGAAGCAGATAGACAATCTCAAGAGTTGGGTATACAGAGC	38051	
Qy	769	GAGTTCCAAGTTTAGTCCATCTCTCATCTTTGTAGCGATTGCAAGCTTATATGTGGGCTG	828	
Db	38052	AGTCTGAGATCAATCTCTCTTTCTACCTTCGTGGTGACTCTAGCTTTCATTTGGGTAA	38111	
Qy	829	GCATAACGAAATCATTCACAGCAGATGAAGACCAAGACACAGGAGTGCATTT-----TT	882	
Db	38112	GCTTGATTAAACACACTTGTTCAGACAGTGAACAGAGGCCCAAGCAAGCTCTTTCACT	38171	
Qy	883	TCTTTGATTCGGTTCGATCTAAGGCCAGATTTAGA-----TCCGCGGTTCTCTGAAATTT	936	
Db	38172	TGATGATCAACGTGGATTGCAAGGAATCGTCTCAAGTACACACAACCAATACCAACAACAT	38231	
Qy	937	ACTTCGGGAAGTCTTATCGTACCGCTGCCGAGAAATCCGCGCGGAGCTGGTGGAG	996	
Db	38232	ACTTTGGCACTGATGGCTCCTGATCGTATCAGTCAAGAAACAGCATTTGCTAGGAG	38291	
Qy	997	AGAAAGGGGTGTTTCTTGCGAGCTGAGGTAAATCGCGGGGGAGATAAAAAAAAGGATCAACG	1056	

	I I
Db 38292	AAAAATGGCTCATGGCTGCTTTCAGATGCACATCAGCGAGAAATCAAAAGATATGTATCAA 38351
Qy 1057	ACAAGAGAATAATTAGAAAACSGTGGGAGAANTGGTCGCCGAGATTGCTGAAGCGTTGTCAGA 1116
Db 38352	CGGA---TCTGTTGAAGACACCACCAAGATGGGACAAGGAGTAGTGTAATATGGGTCATGT 38408
Qy 1117	AATCATATT---TTTTCCGTGCGAGGATCGAGCAAGCTAGACTTTTACGGTGCAGATTTTTG 1173
Db 38409	CCCATTATCCAACCTCAATTTCAGGAGCTCCGAAATTCGGACTGTATCACATGATTTTTG 38468
Qy 1174	GATGGGGAAGCGGAGAAACACGAAATATTTCGATTGATGGGAGAGAAATATGCAATGA 1233
Db 38469	GATTAGGAAAGCCCTGC AAAATGGAGATAGTCACATAGAACACAG--GTGGTCTTATFCG 38525
Qy 1234	CGCTTTGTAAAGCAGGAGTATTTCGAAGCAGGATTCGAGGTTTGCTTTGGCTCCTAAGG 1293
Db 38526	CATTCTCAGATCCAGACAGACGCCACCAATGAGTTGAGATTGGAATACCACTAGAGAAGA 38585
Qy 1294	ACAAATGGATGCTTTTCTGCTTATTTTTCACCTGGGAATTAATGGTTTAATAATGTA 1351
Db 38586	AGAAATGGATGTTGTTTTCACCTGCTTACAAAAGGGGATCAAGAAATTCGCAACGTA 38643
RESULT 10	
AC003027	AC003027 119914 bp DNA PLN 30-JAN-1999
LOCUS	Arabidopsis thaliana chromosome I BAC F21M11 genomic sequence,
DEFINITION	complete sequence.
ACCESSION	AC003027
VERSION	AC003027.1 GI:4079614
KEYWORDS	HTG.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
AUTHORS	Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
	Brassicaceae; Brassicaceae; Arabidopsi
	1 (bases 1 to 119914)
	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
	Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
	Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
	Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.S.,
	Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 119914)
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
	Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y.,
	Oj,O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vyotskaia,V.,
	Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
TITLE	Direct Submission
JOURNAL	Submitted (22-OCT-1997) Biochemistry, Stanford University/DNA
	Sequencing and Technology Center, 855 California Avenue, Palo Alto,
	CA 94304, USA
REFERENCE	3 (bases 1 to 119914)
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
	Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
	Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
	Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
	Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
TITLE	Direct Submission
JOURNAL	Submitted (30-DEC-1998) DNA Sequencing and Technology Center,
	Stanford University, 855 California Avenue, Palo Alto, CA 94304,
	USA
REFERENCE	4 (bases 1 to 119914)
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
	Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
	Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
	Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
	Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
TITLE	Direct Submission
JOURNAL	Submitted (30-JAN-1999) DNA Sequencing and Technology Center,
	Stanford University, 855 California Avenue, Palo Alto, CA 94304,
	USA

COMMENT

On Dec 30, 1998 this sequence version replaced gi:2734094.
 Bases 1-9262 of clone F21M11 overlap with bases 68998-78259 of
 'TAMU' BAC clone F20D22 (AC002411) and bases 119525-119914 of clone
 F21M11 overlap with bases 1-389 of 'TAMU' BAC clone F21B7
 (AC002560).

e-mail for correspondence: arab@sequence.stanford.edu
 Genes with similarity to proteins in the databases are described as
 'putative', '-like' or 'similar to'. Genes that have EST
 similarity but no significant protein similarity are described as
 'unknown proteins'. Genes that are annotated based only on gene
 prediction software are described as 'hypothetical proteins'.
 The software programs used to predict genes include: Grail
 (Informatics Group, Oak Ridge National Laboratory,
<http://compbio.ornl.gov/section/index.html>), GENSCAN (Chris Burge,
<http://genomic.stanford.edu/~chris/GENSCANW.html>), Fexa (V. Solov'yev
 & A. Salanov, Sanger Centre, <http://genomic.sanger.ac.uk/>), and
 NetPlantGene (S.F.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, <http://www.cbs.dtu.dk/NetPlantGene.html>).

FEATURES

source

1. 119914

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="1"

/clone="F21M11"

1. 7000

misc_feature

/note="Proteins in this region are annotated in the F21M11

entry, AC002411."

misc_feature

1. 9262

/note="Overlap with bases 68998-78259 of 'IGF' BAC clone

F21M11."

7684. 10913

/gene="F21M11.1"

join(7684. 7866.7965. 8027.8122. 8160.8250. 8381,

8458. 8728.8875. 8993.10003.10256.10344. 10410,

10501. 10551.10685. 10781.10885. 10913)

/gene="F21M11.1"

/note="Hypothetical protein"

/codon_start=1

/protein_id="AAD10665.1"

/db_xref="GI:4204284"

/translation="MAPNLRIKACDAMKLLGISETKYTRAFRLKLLKTYENNWDFFIE
 DAYKVLDAIFDAOASTKNEKKKEKKKESKSVATSGRRKKAPELPQDDE
 DMDDEDFPLRLRRGRASSSSSSNNEDLKTQPEEDEDGVTLEPLPKRY
 VRRGERGLAMTVNNASPSRSSRLMEPEVPMVLLPAHPMTKYSEASALVIND
 EPNTDRKPVISDTGNCAPMLEMGKSNHVEQWDEWKDILNDTAMDVSPSSAIGES
 SEHKVAASVELASSTSEAKICLSEAPATGETNLHLPMSMEDLRAMEKCLKSYKI
 VHPFESVLGPMKMCSCYIDLAKNSTQLLETVCVDMKAGDESGAVGISMPLVVVP
 ECEISDGWKAISNKDITAGEENVEIPWYNEINEKVPFRYMPHSFVQDAPVIFS
 LSFSDQSCSTCIEDCLASEMNCAGVNDGFAITLDGLLKEEFLEARISEARDQ
 RKQVLFCECEPLERAKKVEILEPCGHLKRGAIKECWKCGCTKRCGNRVVORGHN
 KLOVFFPNKGWGLRLTEKLPKCAFICEVIGETLITPELYQSFEDKPTLVILDAAH
 WGSERLEGKALCLDGMFYGNISRFNLHRCLDANLIEIPVQVETPQHYHLAFTT
 RDIAMELAWDYGIDFNMDLMKPFDCGSGFCRKNRKRSTMQILNKA"
 join(13033. 13488.13717. 13926.14139. 14514)
 /gene="F21M11.2"

mRNA

13033. 14514

/gene="F21M11.2"

join(13061. 13488.13717. 13926.14139. 14316)

/gene="F21M11.2"

/note="Similar to acid phosphatase; Location of ESTs

110C2T7, gb|T42036, and 110C2XP, gb|A1100245"

/codon_start=1

/protein_id="AAD10666.1"

/db_xref="GI:4204285"

/translation="MDRTMFLSLTIALLVGVVSAGDWNILNQLRGSSSSQNGIVS
 KGKTDLKGCSERINVRHNRKFDVPOECVSHIKDYMTSSQYKDDVARTVDEVI
 LHFGMCCSKCDGMDAWIFDIDTLLSTIPYHKNGFFGEGKLNSTGKDWTKKK
 AVPHPMKKLYHDIRGKIFLITSSRKEYLRSATVDNLQIAGYGNLMLRGLEQ
 QKEYQVYKSEKRWKLSLGYRVMGVQDSSFAGCPLPRTFKPLNSIYVVA"
 complement(join(14004. 14312.14565. 14641.14787. 15831,
 15906. 16006.16327. 16362))

mRNA

15906. 16006.16327. 16362)

/gene="F21M11.3"

gene

CDS

complement(14004. 16362)
 /gene="F21M11.3"
 complement(join(14567. 14641.14787. 15831.15906. 16006,
 16327. 16362))
 /gene="F21M11.3"
 /note="Unknown protein; Location of ESTs 40C3TT,
 gb|AA728590 and 40C3T7, gb|T04573"
 /codon_start=1

/db_xref="GI:4204286"

/translation="MLDQISGNTHKSLVETAEPHLNDRVHIVEIPKASVPIPIE
 ICDAEEKCPSTISRRKRVFDSKVITYHVHVSSEVSEELSEENEESEKRSUKSSK
 TDQIIIEVASNSGSIYRHNRYKNCRESDDIEDEDFDCSDLDDEEYYSVDFGSE
 DLHNPTKEVYTDIGDKTEIDSKLRSNETVRDGHYDQGQVNLNENVTOWKSAK
 SKRGTOKQSKNSNPIDAQEERKSDSSFGTDPOIDITLSVKPKCRIRPKKLNRNO
 LAVDASLSTWLSSESSECSNSAMTYLTPEKLSKTSYKSKPLRINHDDRPVLCALTL
 EDKQFASPTPRKSPDEPTPIIGTVGGYMGNRKAIDCGSASSFKGIPNTSSKY
 REDKSVNHSHTPFPEARLEKALNNIDK"
 complement(16835. 20238)

gene

CDS

/gene="F21M11.4"
 complement(join(16835. 17185.17274. 17392.17491. 17656,
 17795. 17885.17982. 18079.18175. 18361.18504. 18604,
 18705. 19049.19134. 19349.19439. 19791.19862. 19967,
 20188. 20238))
 /gene="F21M11.4"
 /note="Hypothetical protein"
 /codon_start=1

/protein_id="AAD10668.1"

/db_xref="GI:4204288"

/translation="MGSSSPPEARAAQVPSMILFLEICTVHVHYNRRKLNDRVLSA
 NLNIPKRVTRDRFMESVSIYKSLNAAVSVHLPLQIPNDCYNKDALNNSPKH
 GESEDSMTDKDVSKRSGTDSRSGSLPTSESDPRPKHQDWTKQLSDHLLYE
 FESEYDAANTPESYTEQAANKNVDRITASEQPSNAARKRICGDSFIOESSPNKPTQD
 TLLRLMESLSDOPTDYVKAQNHOMVYFSEFVIYSFVLCRLSNLRVCVILFCVQQLPK
 SHFEQDSKRKRITASDAMENHLKVPKRENNLQKADIDCNGKCSANSDDLSEKIS
 KALEQTSNITICGFCQSAVSEATGEMLHYSRGPVDDGDDIFRSNVIHVSACIEWA
 POVYEGDTVKNLKAELARGMKIKCTCKSLKGAALGCFVKSCRRSYHVPCEATSRG
 WYDEDFLLCPAHSSVFPNEKSGHRVSRAPLPKINPAELCSLEQTPAFTKELVKGW
 SALSCKDKLMESLAVRFNATISRYWNPSVTHVIASTDEKGACTRTKLVLMLGKGM
 IINAAWMSKASQSPVDERPFEIQIDTQCGDGPKTARLRAETNKPKLEGLKPYFF
 GDFYKGYKDLONLKVAGGTILNTEDELGAESSNNVNDORSSIVYINIDPPHGICAL
 GEEVTIIWQRANDAEALASQTSGLVGHWTWLESIAIGYKLPKHPVIG"
 20752. 24647

gene

CDS

/gene="F21M11.5"
 join(20752. 20994.21362. 21497.21596. 21740.21825. 21994,
 22102. 22178.22542. 22758.22920. 23064.23343. 23442,
 23599. 23693.24021. 24072.24227. 24298.24441. 24647)
 /gene="F21M11.5"
 /note="Hypothetical protein"
 /codon_start=1

/protein_id="AAD10668.1"

/db_xref="GI:4204287"

/translation="MGANSKSVTASFVIAVFFLICGRTAVEDETEFEHGDYSKLSGI
 IIFGASTQLRAWSLDCPYTPDFNLDLWLDLTKLLSAVNCWFKCMVLDPNQYD
 HPCKSRPDSGLSAITELDPGYITGLPSTYMKWKWCVFEGIEANAIVAPYDMRLS
 PTKLERPDLYFHKLLFTFETALKRGPSIVFAHSMGNVFRYLEWLRLEIAEPKHYL
 KWLDDH'THAYFAGCAPLLGSVEATKSTVGLPGVSEGTARLLSFSASSLWLMPE
 SKNCKGDNTEFTHESGCAAKKRVYHCDDEEYQSKYSGMPTNLIETIPSTARELA
 DGTFLKAIEDYDPSDKRMLHQLKKYVFFVIRNIAHRSSLAGFLYIHDDPVENPLTPW
 ERPPKRVFCIYGAHLKATEVGYFAPSGKPYPDNWLITDIYIEGEGSLVSRSGVVDG
 NAGPTGDETVPVHLSWCKNWLIGPKVNTMAPQILIGIKQKQPEHGDGSDVHVELND
 HEHGSIDIANNTRKAPRVKITYFEDSESIPIGKRTAVWELDKSGY"
 27777. 28734

mRNA

gene

27777. 28734

/gene="F21M11.6"

28007. 28465

/gene="F21M11.6"

/note="Unknown protein; Location of ESTs 203124T7,
 gb|H76794 and 203124XP, gb|AA605510"

/codon_start=1

/protein_id="AAD10670.1"

/db_xref="GI:4204289"

gene	CDS
/translation="MNTKTWRLPPRRVLTADKKERDAFISSVTDNPPETAKFPSPDE KLVPSPVNPSPKSSKSTAAAFPIGYNQLMLAGYLSHEYLTQGTIFGQWNNQRAQAESS KIKFSPHVEAECECEPKRRRYRQNLNLRSDGAQLPGIVNPAQLARFLKL" complement(29264..32033) /gene="F21M11.7" complement(29264..32033) complement(join(29264..31015,31414,31484..31587, 31782..32033))	

Query Match	6.7%	Score 99.6;	DB 12;	Length 11914;
Best Local Similarity	47.3%	Pred. No. 1.8e-13;		
Matches 657; Conservative	0;	Mismatches 684;	Indels 47;	Gaps 10;

Qy	4	TGATCGAAACGTGTAGAGTTGGGGCGCGCGGACTCGGTGGCGAGCAATCGCTGCCGC	63
Db	45957	TCATCGAGACTGTACATCTCTCCCCAAAGGCGACGGTTTCCATCAACCACTCTTCCCTC	46016
Qy	64	TCATATTTTTCGACATGACGTGGCTGCTATTTTCATCCCATGCTTTCAGCTCTCTCTACG	123
Db	46017	TCACCTTCTTCGATGCCCTGGCTCTCTCCACACTCGCGGATCTCTCTCTCTCTCTCT	46076
Qy	124	AATTCCCTTGTTCGAAGCAACATTTTTCAGATCCATCGTTCGTTCCAAAACACTCAAA	183
Db	46077	CTTACCAAAAACCTCAACTGAATCTTTTCCCTCCAAAGACTTTGTACCCCAACGCT	46136
Qy	184	TCTCTAAACACTCTCATACACTTCTCCCTCTCTCATGCAATTAATCTACGCTTCATCC	243
Db	46137	TCTCCATCACTCTCCACATTTCTTCCCTTACGCCGGTAAACTGATTATCCCGGCTCG	46196
Qy	244	CGGAGAAAATGCCGAGTTTCGGTATCTATCCGGGGACTCGGTTTCTTACCACTCCGAG	303
Db	46197	CCGACCCTCCATATTTACACTACAACGATGCCAAGACTCTCTTGTTCACCGTAGCAG	46256
Qy	304	AATCTACGGA---CGACTTCGATGATCTGTCGGAATCTGTCAGAACTCTCCCGTTAGGC	360
Db	46257	AGTCTACTGAACACGACTCTCGAACACTGAAGTCTGATTCGCCCTTAAAGATATCATG	46316
Qy	361	TCTACAACTTTGTCCCTAAATTGCGGCCACTTGTCTGAAGAAATCGGATAGAA---AACT	417
Db	46317	TGCATGGGGCTTTCGCCCAAGTTACTCTCTCACGCTCTCCCGAAGGAATTCAGATGC	46376
Qy	418	TCCAAGTTTTCGCGTGCAGTGACTCTTTTCCAGGCCGAGGGCTCGGTATTTGGAATAG	477
Db	46377	GACCAATAATGGCTATGCAAGTCAACACTCTTCCCTGGAGCTGGAATCTGTATAGCAACT	46436
Qy	478	CAACGCATCACACGGTTAGCGACGCCCGCTTTCGCGGTTTATACGGCTTGTGCTTT	537
Db	46437	CAGCTACACATGTTGTAGCAGATGGAGTCACCTTCAGTCAATTTTCATGAAGTATTTGGAT	46496
Qy	538	CAATGAGCAACACATTTGAAATCAAGATGAAGATGAAGATTTTAAATCTTTGCTTCG-	591
Db	46497	CATTGACCAATTCAGCGGTTAAGATCCGCCACGGTTCTTCTACCTCTCTCGCTATTC	46556
Qy	592	-CAGTTTTTCGATAGATCCGTCATAAAATATCCGACGAAATTTTGACTCCATTTATTGGAGA	650
Db	46557	ACAGCTCGAGAAACATGATCAAGGACCCAGCGAGGTAGGCGCAGGACATTTTGAGCGGT	46616
Qy	651	AACCGGCTAAATTTCTTTGCAATCTCTGTCATCCCTCAITTTACCGAGGACCGCATTCGA	710
Db	46617	TTTGGAGCCAAAACCTCTGCAAAACACAGTTCACATGTTACGCCCTGA-GAACATGGT	46675
Qy	711	ACCACGTTTCGTTTTTCACCCCAATCCAAAATTAAGAAATTTGAAGGGTTGGATTCAGTCC	770
Db	46676	GCCACTTTTACATTTGAGCCGGAACGATAGATATCTAAAAAGTTGGTTACAGAGCAG	46735
Qy	771	GTTCCAGATTGTAGTCCATCTCTATCTTTTGTACGGATTTGACGCTTATATGTGGGCTGGC	830
Db	46736	TCTGAGAAATCAATCTCCTGTTTCTACCTTCGTGGTGACTCTAGCTTTCAATTTGGTGA	46795
Qy	831	ATAACGAAATCATTT-----CCCC-----CACACGATGAAGACCAAGCAACAGGAT	875
Db	46796	TTGATCAAGACACTTTGTACAAGACAGTGAACAAGGCCAACAGGAGAACAGAGATGAA	46855
Qy	876	GCATTTT---TTCTTTGATTTCCGGTCTGATCTAAGGCCAGGATTTAGA-----TCCGCCGGTT	926

DADAIEANDDKVSBILKEGAMIFSVAGSTQFDVYGSDFGWRGPKKKVEIVSDRTQ
AISAERRDGGGVGVVLEKQOMEFESVFADGLKNDLV"
BASE COUNT 464 a 344 c 354 g 486 t
ORIGIN

Query Match	5.2%; Score 76.8; DB 9; Length 1648;
Best Local Similarity	50.9%; Pred. No. 3.4e-08;
Matches 342; Conservative 0; Mismatches 312; Indels 18; Gaps 6;	
QY 812	ACGTTATATGTGGGTGCGATACCAATATTCACAGCATGAAGACCAAGACAACA 871
DB 990	AGCTAAATAGCAAAAGCAAACTGAAGCTGACAGCATATGATGAATATAA 1049
QY 872	GGATGCATTTTCTTGATTCGGTGCATCTAAGGCCACGATTTAGATCCGCCGTTCCCTGA 931
DB 1050	AAATATTATTGGGATTCACCTGCGGATATATAGGACCGTTTGGATCTCTCAATCCACT 1109
QY 932	AAATTACTTCGGGAACCTGCTTATCGTACGCGTCCGCGGAATCGCGCGGACGCTCGT 991
DB 1110	TAATTTATTTTGTAACTGC--AATGGAGACATTTGTGAGACTGCACCAAGCAAGTGATTTTC 1167
QY 992	GGGAGAGAAAGGGTGTCTGCGACGTCGAGTAACTCCGCGGAGATAAAAAAGGAT 1051
DB 1168	GTTCAAGAAATGGGGTGTCT--TTTGTTCAGAGATGTTAAGTGATATGGTCAAAAGGAT 1226
QY 1052	CAACGACAGAGAATATTAGAAACGGTGGAGAAATGGTCGCGGAGATTCGTTAAAGCGTT 1111
DB 1227	CGATG--CGGATGCCATTGAAGCCAAATGATGATAAGTTTCAGAAATATTGGAATCTTG 1284
QY 1112	GCAGAA-----ATCATATTTTTCGTCGAGATCGACAAAGCTAGATCTTTACGGTGCA 1166
DB 1285	AAAGAAGGAGCAATGATTTTCTGTGCTGCTCGACCCCAATTTGATGTTTACGGGTGCG 1344
QY 1167	GATTTTGGATGGGCAAGCGGAGAAAGCAAGAAATATTGTCGATTTGATGGGAGAAATAT 1226
DB 1345	GATTTTCGGGTGGGGAGGCCCAAGAAAGTGGAGATGTTGTTCAATAGATAG--GACACAA 1401
QY 1227	GCAATGACGCTTTGTAAGCCAGGATTTGCAAGGAGGATTTGGAGGTTTGGTGTCTTTG 1286
DB 1402	GCCATCTCTTTGGCAGAGAGAGATGGAGGCGCGTGGAGTTGGAGTGTCTTTTA 1461
QY 1287	CCTAAGGACAAATGGATGCTTTTCTGCTCTATTTTTCACATGGGAATTAATGGTTAATA 1346
DB 1462	GAGAAGCAACAATGAGGTTTTCGAATCTGATTTGCTGCTGAGTGA--AAATG 1516
QY 1347	ATGATGTAATTAATACTAATATTATATGTAACATTAATTAAGTGTTCAGTAACTGAA 1406
DB 1517	ATCTGTTTAATTAATGATGATCATCTAAATTTCTCAATATATTATTGGTCATATTCAA 1576
QY 1407	GAATAATCCCTATTATATATTATGATTTGGTTCAATAAAGTGTAAAGCCTCTTTGAAAA 1466
DB 1577	AAGAAATAATATTTCGGGATTTTGTGACCACCAATAATAATACTCTTTTGAAGAAA 1636
QY 1467	AAAAAAAAAAAA 1478
DB 1637	AAAAAAAAAAAA 1648

RESULT 12
AB028618 85690 bp DNA PLN 27-DEC-2000
LOCUS Arabidopsis thaliana genomic DNA, chromosome 3, p1 clone:MOD1.
DEFINITION Arabidopsis thaliana
ACCESSION AB028618 BA000014
VERSION AB028618.1 GI:5041971
KEYWORDS
SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui p1
clone:MOD1.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (sites)

AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL

Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S.
Structural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
and BAC clones
DNA Res. 7 (3), 217-221 (2000)
20363099
2 (bases 1 to 85690)
Kaneko,T., Katoh,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S.
Direct Submission
Submitted (09-JUN-1999) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)
Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/aggd_graph.cgi?c=MOD1
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
Netgene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremlin1.zoool.iastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is T13J10 and the 3' clone is T26G12.

FEATURES
source

1. .85690
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MOD1"
/clone_lib="Mitsui p1"
2123. .8637
/note="contains similarity to TNP2-like transposon
protein
gene_id:MOD1.1"
/codon_start=1
/pseudo
/evidence=not_experimental
join(10671..10844,10921..11208)
/note="contains similarity to unknown protein
emb|CAA16722.1
gene_id:MOD1.2"
/codon_start=1
/pseudo
/evidence=not_experimental
15142. .16497
/note="gene_id:MOD1.3"
/codon_start=1
/evidence=not_experimental
/product="anthocyanin 5-aromatic
acyltransferase/benzoyltransferase-like protein"
/protein_id="BAB02518.1"
/db_xref="GI:11994477"
translation="MTLHVLETARVTPDYSVINSANLKLPLTFEFLPWLFPVKR
VFVELTESTRDFHSLILPKLKSLSLILRNLYPLTGHTWEENEKPSIISBNGV
VLVTIAESDADFSLHSGYQRPSELHALPKLPVSDSATAFSIQTTLFPNQFSIG
RSAAHVAIDKGTSSFTKAWAQICKQLQSPENLTPSYDRSLIKYPTLDERMIELV
VKLEDQNTIRSLTSLPSSKGLDVLATVLSRADLERLEQVKNVSPSLHSTFVI
AYAWNTCFVKARGKNDRSVSLFVGDFDRDLPKLPQTVFGNCMPVGCYNKAAE
FMEKGFVTAEEILSDVGLSSRKIIETIADTFVEGFSQSWSTQFTIAGSTRIGVY
EADPGWRPVRKVIDIVSDQGEATAMARRDESQGVETGMCLKTKTMDSVSVSFFNNGLH
S"
complement(18639..19994)

CDS

Db 15557 ATGACTCAGCGA-----CCGCTTTCTCCATCAAAATCATCGTTGTTTCCTA 15601
Qy 454 GCCGAGGCGTCGGTATTGGAATAGCAAGCATCAACACGTTAGCGACGCCCGCTGCTTC 513
Db 15602 ACCAAGGATTCTCCATAGTGTTCAGCACACCATGCTGTTTAGCGGGAAGCTCAA 15661
Qy 514 TCGCGTTTATACGCGCTGGCTTCAATGAGCAACA--CATTTGAAATGAAGATGAAGA 571
Db 15662 GCACCTTTTATCAAGCTTGGGCTCAAACTCTGCAACAAGAACTTCAAGACATGCCGGAGA 15721
Qy 572 TGAAGAAATTAATCTTTGCGAGTTTTCGATAGATCCGTCATAAAATATCCGACCAAT 631
Db 15722 ATCTAACCTCGCTTTATGACCGGCTTTTGATCAAAATATCCGACCTATCTTGATGAAAGA 15781
Qy 632 TGACTCCATTTATTGGA-----GAAACGCGCTAAATTTCTTCGAATCTCGTCATCCC 686
Db 15782 TGATAGAGTAGTGAGATCTCTCAAGAGGACCAACCAATCAGAGCTTGATCTCGC 15841
Qy 687 TC-----ATTACGAGCGACCGCATTCGAACACGCTTGGTTTACCCCAATCCA 735
Db 15842 TTCCCTCTAGTAAACTCGGAGACGATGTCGTGTAGCCACGCTAGTCTCTCTCGGGCTG 15901
Qy 736 AAATTAAGAAATGGAAGGTTGATTCAGTCCAGAGTTCGAAGTTAGTCCATCTCTCAT 795
Db 15902 ATATTGAGAGACTTAGGGAACAAGTCAAGAAGCTGTACCCGAGCC---TCCACTTGTCTGA 15958
Qy 796 CTTTGTAGCGATTGCAGCTTATATGTGGCTGGCATAAACGAATCATTCACAGCAGATG 855
Db 15959 CTTTGTCTATTGCTTACCGCTACCGCTGGACTTGTTCGTAAGCCAGCTGGAGTAATA 16018
Qy 856 AAGACAAGACAACGAGATGCATTTTCTGATTCGGTGCATCTAAGGCCACGATTAG 915
Db 16019 AAGACAGA-----TCAGTGAAGTCTCTTGTTCGTAGGAGATTTACAGACCGGTTAG 16069
Qy 916 ATCCGCCGGTTCCTCAAAATTTACTTCGGGAAGTCTTATCTGACCGCTGCCGAGATGC 975
Db 16070 ATCCGAAGCTCCCGGGGACTTACTTCGGAATCTGTATGATTCGGTGGGCTGTTATAACC 16129
Qy 976 GCGCGCAGAGCTGGTGGGAGAGAAAGGGTGTTCCTGGCAGCTGAGGTAATCGCGGGCG 1035
Db 16130 GTAAGCGCGGAGCTTTATGAAGAGAAGGNTTGTGACCGCGGCTGAGATATTACCG 16189
Qy 1036 AGATAAAAAAAGGATCAACGACAGAGAATATTAGAAACCGGTGGAGAAATGGTCGCGG 1095
Db 16190 ATTTGGTGAAGAGGTTGAGTTCAAGAAAGATAGACACATTCGCGACACTTTCGTGGAG 16249
Qy 1096 AGATTCGTAAGCGTTGAGAAATCATATTTT---TCGGTGGCAGGATCGAGCAGCTAG 1152
Db 16250 GATTTAGTTTTCAGAGTTGGAGTACACAGTTTGGGACGATAGCCGGGTTCGACCGGTTGG 16309
Qy 1153 ATCTTTACGGTGCAGATTTTGGATGGGGAAGGCGAGAGAAACGAAGAAATATTGTCGATTG 1212
Db 16310 GAGTATACAGCGGATTTCCGGTGGGGAAGACCCGTTAAAGTTGATATTGCTCCATTG 16369
Qy 1213 ATGGGGAGAAATATGCAATGACGCTTTGTAAGCGAGGATTTTCAAGGAGGATTTGGAGG 1272
Db 16370 A---CCAGGAGAAAGCGATCCGAATGGCTGAGAGACGATGATGAGTGCAGTGGCGTTGAGA 16426
Qy 1273 TTTGCTGTCTTTGCTTAAGACAAAATGGATGCTTTTTCGCTGCTTATTTTTCACATGGGAA 1332
Db 16427 TTGAATGTGTGTTGAAAAGACTGMAATGGATTCGGTGTCTCTCTTTTCAACAATGGTT 16486
Qy 1333 TTAATGGTTAA 1343
Db 16487 TACATAGCTAA 16497

RESULT 13
AC069470
LOCUS Arabidopsis thaliana chromosome 1 clone IGF-F7F7. *** SEQUENCING IN
DEFINITION PROGRESS ***, 4 unordered pieces.
ACCESSION AC069470

AC069470.10 GI:12320593
HTG: HTGS_PHASE1.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 136047)
AUTHORS Lin,X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Ronning,C.M.,
White,O., Nierman,W.C. and Fraser,C.M.
TITLE Arabidopsids thaliana 'IGF' BAC 'F7F7' genomic sequence near marker
'ml72'
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 136047)
AUTHORS Town,C.D. and Kaul,S.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
On Jan 19, 2001 this sequence version replaced gi:12280790.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 100163: contig of 100163 bp in length
* 100164 100213: gap of unknown length
* 100214 112623: contig of 12410 bp in length
* 112624 112673: gap of unknown length
* 112674 116366: contig of 3963 bp in length
* 116637 116686: gap of unknown length
* 116687 136047: contig of 19361 bp in length.
FEATURES
Location/Qualifiers
1..136047
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="IGF-F7F7"
BASE COUNT 46028 a 22872 c 22213 g 44783 t 151 others
ORIGIN
Query Match 5.1%; Score 75.4; DB 74; Length 136047;
Best Local Similarity 46.8%; Pred. No. 1.5e-07;
Matches 623; Conservative 0; Mismatches 651; Indels 57; Gaps 10;
Qy 37 ACTCGGTGGCGAGCAATCGGTGCGCTCACATTTCTTCGACATGACGTGGTGCATTTTC 96
Db 90101 ACTCAGCCAACCTTGCAAAACTTCTTTAACTTTCTTCGACCTCCATGCTACTATTCC 90160
Qy 97 ATCCCATGCTTCAGTCTCTTCTTACGAATTCCTTTGCCAAGCAACATTTTTCAGAA 156
Db 90161 ACCCAGTCAAAAGAGCTCTTCTTACGAGCTCACCGAGCTACTCGTGACCATTTCCACT 90220
Qy 157 CC---ATCGTTTCCAAAACATCTCTCTCTAAACCTCTATACACACTTCTTCCCTC 213
Db 90221 CCAATTATCTCCCTAAAGCTCAAGGACTCTCTGCTCTATCTCGTAACTATCTCCCTC 90280
Qy 214 TCTCATGCAATTTAATCTTACCTTCATCCCGGAGAAATGCCGAGTTTCGGTATCTAT 273
Db 90281 TCATGGCCACATCATTCTGGGAGCCAAACGAGCAATATCGTCTCGGAAA 90340
Qy 274 CCGGGGAGCTCGGTTTCTTTTCACCATCGCAGAATCTAGCGACGACTTCGATGATCTCGTCG 333
Db 90341 ACGGGGTC---GTTTGGTGACTATAGCCGAGAGTGACGCTGACTTTTCTCTCTTTCCG 90397
Qy 334 GAAATCGTCCAGATCTCCCGTTAGGCTCTACAATTTGTCCCTAAATTTGCCGCCCATTTG 393
Db 90398 GTTATGGACAGCGCTCTTTGTGACAGTTTACATGCTGTGTTCCCAAGTTACCGGTTTCTG 90457

Db 28555 GTACGATCTCGCTCGAGTTGACTCGAGAGATATAGAGAACTTAGGAGCAAGCCCAAG 28496
Qy 765 TCCAGAGTTCCAAAGTTTAGTCCATCTCTCATCTTTTAGCGATTGCGACTTATATGCG 824
Db 28495 AGGGAGTCAGCTCGCTCATCATCATGAGTTCACCTGTCAACGTTGTTGCCAACGG 28436
Qy 825 GCTGGCATACCAAGATCATTCACAGAGATGAAGACCAAGACAAGAGAGATGCAATTTTC 884
Db 28435 TATCTTTGACATGCTTGTGTAAGACGGCTGGAGAGATGAGAACAGACCGTCCGTTTC 28376
Qy 885 TTGATTCCGGTCGATCTAAGCCACGATAGATCCCGCGTTCTTGAAATTTACTTCCGG 944
Db 28375 ATGATGCTGCTGATTTTCAGGAACCGTTAGATCCACGGTTCCCGAGATGATTTCCGG 28316
Qy 945 AACTGCTTATCG- ---TAGCGCTCCGAGAAATGCGCGGAGAGCTGGTGGGAGAA 1000
Db 28315 AACTCGGTGTTCCCGATAGTTGCTTTGGTTACAGGCGAATGTTTTGGGAGAAAT 28256
Qy 1001 AGGGTGTGTTTCGAGCTGAGGTAAATCCCGCGGAGATAAAAAAGGATCAACGCAA 1060
Db 28255 GGGTTGTTAATATGCTCGAAATTTCTAAGTGATTCGGTTAGAAGTATTGGTTACGAA 28196
Qy 1061 GAGATATTAGAAACGGTGGAGAAATGTCGCCGAGATTCGTAAGCGTTGCAGAAATC 1120
Db 28195 TTGGAACAAATTTGCGAATTTGCAATTTACATTAATGGAACAAAGAGTGTGAACACAGGTACACAG 28136
Qy 1121 ATATTTTTCGGTGGCAGGATCGAGCAAGCTAGATCTTTACGGTGCAGATTTTGGATGGG 1180
Db 28135 AT-TGGATCCATTGCGGTTCGAACCACTTTGGGTATACGGTTCAGATTTTCGGTGGG 28077
Qy 1181 GAAGCGGAGAAAGCAAGAAATATTTCGATGATGGGAGAAATATGCAATGACGCTTTG 1240
Db 28076 AAACCATGTAAACAGTGAAGTTCGCTCCATTGACCGA- ---AACGAGCGTTTTCGATGTC 28020
Qy 1241 TAAACCCAGGATTCGAAGAGGATTCGAGGTTTGGCTTGTCTTTCCTTAAGGACAAAT 1300
Db 28019 GGAGAGGAGGATGAACCCGAGGTTGGAGATCGGCTTGTGTTGAAGAAATGTGAGAT 27960
Qy 1301 GGATGCTTTGCTGCTATTTTTACTGGG- -AATTAATGGTTAATAATGATGATTAAT 1358
Db 27959 GGACATATTATTATCTGTTCAAAATGGTTTAGAAATGGTCTTTAGGGTCTTTGATT 27900
Qy 1359 AACTAATATTATTATGTAACAAATTAATTAAGTGT 1394
Db 27899 CTAATGTTTGAAGATGTAATTTTTTTTTTCTTTT 27864

RESULT 15
166494/c
LOCUS I66494 7218 bp DNA PAT 28-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION I66494
VERSION I66494.1 GI:2724471
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 7218)
AUTHORS Dörner, F., Schellflinger, F. and Falkner, F. Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
Location/Qualifiers
1. 7218
/organism="unknown"
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN
Query Match 4.4%; Score 65.8; DB 10; Length 7218;
Best Local Similarity 7.5%; Pred. No. 2.1e-05;
Matches 31; Conservative 219; Mismatches 161; Indels 0; Gaps 0;
Qy 862 AAGACAACGAGGATGATTTTTTTTCTTTGATTCGGTGTCTAAGGCCAGGATTAGATCCGC 921

Db 1505 AAAACGGCATGTAGCATCACTGTAAATTACCTATCTATGCAAGTAGTTAAAGAGATAGAA 1446
Qy 922 CGGTTCTCTGAAAATTACTTTCGGGAACCTGTTATCTATCGCTGCGGAGAAATGCGCGCGC 981
Db 1445 GAATTTGGTACRR 1386
Qy 982 GAGAGCTGGTGGGAGAGAAAGGGGTGTTTCTGGCAGCTGAGTAATCGCGCGGAGATAA 1041
Db 1385 RRR 1326
Qy 1042 AAAAAGGATCAACGACAAGAGAAATATTAGAACCGTGGAGAAATGTCGCGCGGAGATTC 1101
Db 1325 RRR 1266
Qy 1102 GTAAACGCTTCAGAAATCATATTTTTCGGTGGCAGGATCGAGCAAGCTAGATCTTTAGG 1161
Db 1265 RRR 1206
Qy 1162 GTGAGATTTTGGATGGGGAAGCGAGAAAGCAAGAAATATTGCGATTGATGGGAGA 1221
Db 1205 RRR 1146
Qy 1222 AATATGCAATGACGCTTTGTAAAGCCAGGATTTTCGAAGGAGGATTTGGAGG 1272
Db 1145 RRR 1095

Search completed: November 5, 2001, 16:50:26
Job time: 10796 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:11:37 ; Search time 445.49 Seconds
(without alignments)
2084.596 Million cell updates/sec

Title: US-08-894-356C-4
Perfect score: 1479
Sequence: 1 CCGTGATCGAACGTGTAGA.....TTGAAAAAAAAAAAAAAAAA 1479

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 730101 seqs, 313950809 residues
Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601.*
1: /SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1478.6	100.0	1479	17	AAT37311 Aromatic acyl tran
2	760.6	51.4	1518	17	AAT37313 Aromatic acyl tran
3	216.8	14.7	1622	17	AAT37309 Aromatic acyl tran
4	190.6	12.9	1508	17	AAT37312 Aromatic acyl tran
5	175.4	11.9	1703	17	AAT37308 Aromatic acyl tran
C 6	94.8	6.4	936	22	AAF58252 Oligonucleotide D1
C 7	94.8	6.4	936	22	AAF58254 Oligonucleotide D1
C 8	94.8	6.4	936	22	AAF58257 Oligonucleotide D1
C 9	94.8	6.4	936	22	AAF58259 Oligonucleotide D2
C 10	94.8	6.4	936	22	AAF58262 Oligonucleotide D2
C 11	94.8	6.4	938	22	AAF58255 Oligonucleotide D1

12	91	6.2	936	22	AAF58252 Oligonucleotide D1
13	91	6.2	936	22	AAF58254 Oligonucleotide D1
14	91	6.2	936	22	AAF58257 Oligonucleotide D1
15	91	6.2	936	22	AAF58259 Oligonucleotide D2
16	91	6.2	936	22	AAF58262 Oligonucleotide D2
17	91	6.2	938	22	AAF58255 Oligonucleotide D1
18	76.8	5.2	1648	21	AAAC64785 Lemon acyl transe
19	72.6	4.9	1379	21	AAAC46852 Arabidopsis thalia
C 20	65.6	4.4	1240	21	AAAC33227 Arabidopsis thalia
C 21	64	4.3	244	22	AAF58238 Arabidopsis thalia
22	62	4.2	1502	21	AAAC49421 Arabidopsis thalia
23	62	4.2	1504	21	AAAC33536 Arabidopsis thalia
24	61.6	4.2	244	22	AAF58238 Oligonucleotide D1
25	45	3.0	1820	20	AAZ00434 Human secreted pro
26	45	3.0	5760	6	AAAN50530 Sequence encoding
27	44.4	3.0	3690	21	AAF15891 Human prostate can
28	42.8	2.9	14066	20	AAAX99556 Nucleic acid seque
29	42.6	2.9	1240	13	AAQ24177 Tox2a gene. Pyemo
30	42.4	2.9	3426	22	AAAC67368 Human NFAR-1 codin
31	42	2.8	580	22	AAAC92918 Human ISF CDNA clo
32	42	2.8	2620	22	AAAC92908 Human immune suppr
33	41.6	2.8	1556	21	AAAC66427 Human secreted pro
34	41.4	2.8	1520	21	AAAC42254 Arabidopsis thalia
35	40.2	2.7	1613	21	AAAC46500 cDNA encoding enzy
36	40	2.7	1569	21	AAAC59282 Human secreted pro
37	39.6	2.7	1082138	21	AAF22305 Arabidopsis thalia
C 38	39.4	2.7	1150	20	AAAX13596 Enterococcus faeca
39	39.4	2.7	1241	19	AAV29479 Nucleic acid encod
40	39	2.6	2401	22	AAAF56506 Arabidopsis CDC27A
41	39	2.6	2434	22	AAAF56505 Arabidopsis CDC27A
42	38.8	2.6	2833	21	AAAF15797 Human prostate can
C 43	38.8	2.6	79122	21	AAAF22294 BAC containing rep
44	38.6	2.6	1217	20	AAZ24836 Human secreted pro
45	38.6	2.6	1583	21	AAAF15676 Human prostate can

ALIGNMENTS

RESULT 1
AAT37311
ID AAT37311 standard; cDNA to mRNA; 1479 BP.
XX AAT37311;
XX AC
XX AAT37311;
DT 06-FEB-1997 (first entry)
XX Aromatic acyl transferase coding sequence.
DE Aromatic acyl transferase; transformation; anthocyanin pigment;
XX plants; acylation; colour; tone; colouration; colour change;
KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;
KW Scenecio cruentus; Lavandula angustifolia; ds.
XX Perilla ocimoides (Clone pSAT208).
OS
XX Key Location/Qualifiers
FH CDS 3..1343
FT /*tag= a
FT /product= Aromatic acyl transferase.
XX WO9625500-A1.
XX PD 22-AUG-1996.
XX PF 16-FEB-1996; 96WO-JP00348.
XX PR 30-JAN-1996; 96JP-0046534.
XX PR 17-FEB-1995; 95JP-0067159.
XX PR 29-JUN-1995; 95JP-0196915.
XX (SUNR) SUNTORY LTD.
PA
XX

PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
 PI Nakao M, Tanaka Y, Yonekura K;
 XX
 DR WPI: 1996-393401/39.
 DR P-PSDB; AA04725.
 XX
 PT DNA coding for aromatic acyl transferase - for transforming plants
 PT which produce anthocyanin pigments and thus altering colour tone,
 PT e.g. of flowers
 XX
 PS Claim 4; Page 65-69; 94pp; Japanese.
 XX
 CC Vectors containing DNA fragments encoding proteins of plant origin
 CC with aromatic acyl transferase activity may be used to transform
 CC plants which produce anthocyanin pigments. The aromatic acyl
 CC transferase acylates the pigments in the flower resulting in colour
 CC tone changes and allowing new colourations to be produced. Six
 CC specific DNA sequences encoding aromatic acyl transferase from
 CC different plants are described in AAT37308-T37313.
 XX
 SQ Sequence 1479 BP; 420 A; 316 C; 331 G; 411 T; 1 other;

Query Match 100.0%; Score 1478.6; DB 17; Length 1479;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1478; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGTGATCGAAACGTTAGAGTTGGCGCGCGCGGACTCGGTGGGAGCAATCGGTGC 60
 DB 1 ccggtatcgaaacggtgtagagttggtggcgccgagctcggcgagcaatcggcgc 60

QY 61 CGCTCACATCTTCGACATGACGTGGCTGCATTTTCATCCCATGCTTCAGCTCCCTTCT 120
 DB 61 cgctcacattcttcgacatgacgtggctgcatttcacatcctcctcctctct 120

QY 121 AGCAATTCCTTGTTCACAGCAACATTTTCAGATTCATCGTTCACAAACTCAACAAAT 180
 DB 121 acgaatcccttgttccaaagcaacattttcagaatccatcgttccaaactcaacaa 180

QY 181 CTCCTCTCTAAAACCTCTCATACACTTCTCCCTCTCTCATGCAATTTAATCTACCCCTTCA 240
 DB 181 ctctctctaaaactctctacactctctccctctctctcgcagatctaatctaccccttca 240

QY 241 CCCCAGAGAAATCGCGAGTTTCGGTATCTATCCGGGACTCGGTTCCTTTCAACCATCG 300
 DB 241 ccccgagaaaatcgcgagtttcggatctatccatccggagactcgtttcttccaccatcg 300

QY 301 CAGAACTAGCGACACATTCGATGATCTCGTCGGGAAATCGTCCAGAACTCCCGTTAGGC 360
 DB 301 cagaactagcgacgacttcgatctcgtcgtggaaatcgtccagaatctcccgcttaggc 360

QY 361 TCTACAACCTTTCTCCCTAAATTCGCGCCCATTTGTCGAAGATCCGATAGAAACTCTTCC 420
 DB 361 tctacaacttttctccctaaatttcgcccattgctcgaagaatccgatagaaaactcttcc 420

QY 421 AAGTTTCCCGCTCGAGGTGACTCTTTTCCAGCGCGAGCGGTTCGGTATTTGGAATAGCAA 480
 DB 421 aagtttcccgctcgaggtgactctttccagcgccgagcgtcggtatttggaatagcaa 480

QY 481 CGCATCACACCGTTAGCGACGCCCGTCTGTTCTCGCGTTTATACGGCTTGGTCTTCAA 540
 DB 481 cgcatcacacggttagcgacgcccgctggttctcgcgctttataaacggttggtcttcaa 540

QY 541 TGAGCAAAACACATTTGAAATGAAGATGAAGATGAAGATTTAAATCTTTCGAGTTTCG 600
 DB 541 tgagcaaacacatttgaaatgaagatgaagatgaagatgaagatgaagatgaagatgaag 600

QY 601 ATGATCCGCTCATAAAATATCCGAGCAATTTGACTCCATTTATTTGAGAAACGCGCTAA 660
 DB 601 atgatacgcgtcataaaaatccgagcaaatttgactccattttatttgagaaacgcgttaa 660

QY 661 AATTTCTTTTGCATCTCGTCACTCCCTCATTTACCGACGCGCATTCGACACCGGTTTCG 720
 DB 661 aatftcttttgcattctcgtcatctccctcatttaccgacgcgcattcgacacgggttcg

DB 661 aatttctcttgcattctcgtcatctccctcattaccgacgagcgaattcgaaaccaagttcg 720
 QY 721 TTTTCACCCCAATCCAAATTAAGAAATTAAGAGTTTGAATTCAGTCCAGAGTTCCAAAGTT 780
 DB 721 tttttcaccccaatcccaaattaagaaatgaaggggttgagttcagttccagaggttccaaagt 780

QY 781 TAGTCCATCTCTCATCTTTTGTAGCATTCGACGTTATATGTGGCTGGCATTAACGAAAT 840
 DB 781 tagtccatctctcatctttttagcgattgcagcttatatgtggcgtgcataacgaaat 840

QY 841 CATTCACAGCAGATGAACCAAGACACAGGAGTGCATTTTCTTGATTCGGGTCTGCATC 900
 DB 841 cattcacagcagatgaagaccaagacaacagagatgcattttcttgattccggtcgcgac 900

QY 901 TAAAGCCACGATTAGATCCGCGGTTCTCGAAATTAATCTCGGGAACCTGCTTATCGTACG 960
 DB 901 taagggccacgattagatccgcggttccctgaaaactactcgggaactgcttatcgtagc 960

QY 961 CGCTCCCGAGAATCGCGCGCGAGAGCTGGTGGGAGAGAAAGGGGTGTTTCTGGCAGGTG 1020
 DB 961 cgctcccgagaatcgcgcgcgagagctggtgggagagaaagggtgttctctggcagctg 1020

QY 1021 AGGTAATCGCGCGGAGATGAAGAAAGATCAACGACAGAGAATATAGAAACGGTGG 1080
 DB 1021 aggtaatcgcgcgagagataaaaaagatcaacgacaagagaatattagaaacggtgg 1080

QY 1081 AGAAATGTCGCGCGAGATTCGTAAGCGTTTCAGAAATCATATTTTTCGGTGGCAGGAT 1140
 DB 1081 agaaatggtcgcgcgagatttcgttaaaagcgttgcagaatacatatttctcgttcggcaggt 1140

QY 1141 CGAGCAAGCTAGATCTTTACGGTGCAGATTTTGGATGGGGAGCGGAGAAAGCAAGAAA 1200
 DB 1141 cgagcaagctagatctttacggtgcagattttggtatggggagagcgagaaagcaagaaa 1200

QY 1201 TATTTCTCATTTGATGGGAGAGAAATATGCAATGACGCTTTGTAAAGCCAGGGATTTGCGAAG 1260
 DB 1201 tattgtcgtattgattgggagaaatgcgaatgacrcctttgttaaagccagggtatttcgaag 1260

QY 1261 GAGGATTCGAGGTTTGTCTTTCCTTAAGGACAAATATGATGATGCTTTTGTCTGCTTATT 1320
 DB 1261 gaggatttcgaggttgcgttgcgttccttaaggagcaaaatggatgcttcttgcgtctatt 1320

QY 1321 TTTTCACTGGGAATTAATGTTTAAATGAATGATGTAATTAATTAATTAATTAATTAACA 1380
 DB 1321 ttctcactgggaattaaatggtttaataaatgatgtaattaaactaatattattatgaaca 1380

QY 1381 ATTAATTAAGTGTTCAGTAAAGTGAAGATATATCCCTATTATATATATATATATATGATTTGGTTC 1440
 DB 1381 attaatgaagtgttgagtaacgtgaagaataatccctattatattattatgatttggttc 1440

QY 1441 AAATAAAGTGAAGGCTCTTTGAAAAAATAAAAAAATAAAAAA 1479
 DB 1441 aaataaagtgaagcctcttgaaaaaataaaaaaataaaaaa 1479

RESULT 2

AAT37313

ID AAT37313 standard; cDNA to mRNA; 1518 BP.

XX

AC AAT37313;

XX 06-FEB-1997 (first entry)

XX

XX Aromatic acyl transferase coding sequence.

XX

XX Aromatic acyl transferase; transformation; anthocyanin pigment;
 KW plants; acylation; colour; tone; colouration; colour change;
 KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;
 KW Scenecio cruentus; Lavandula angustifolia; ds.

XX

OS Lavandula angustifolia (Clone pLAT21).

XX

FH Location/Qualifiers

Db 1199 tcaacatctgcggcaccaggctcgacgctcatttatgtctgttgatccgcacacgcaat 1258
 QY 1149 CTAGATCTTTAGCGGCGAGATTTTGGATGGGGAGCGGAGAACGACGAATATTTGCG 1208
 Db 1259 gtcgatgatttgatgcagatttgggtgggaagccttgaaagcatgaatctgtttca 1318
 QY 1209 ATTGATGGGAGAAATATGCAATGACGCTTTGTAAGCCAGGATTTTCAAGGAGGATTG 1268
 Db 1319 actaatcttcggcaacactaatttgatctctcggtccagaagatttaaaagagcactt 1378
 QY 1269 GAGGTTGCTCTCTTTCCTAAGGACAAATGATGCTTTTCTGCTTATTTTTCACGTG 1328
 Db 1379 gagcttggaattcttctgactgaagaataggtgagcgaatttgcacacattttacgaat 1438
 QY 1329 GCAATTAATGCTTAATAATGATGTAATTAACATAATTAATTAATTAATTAATTA 1388
 Db 1439 ttcataaatagctccatgtgaggagcccttggtaagaaaaagtggtatcaatgtataa 1498
 QY 1389 AGTGTGAGTAACGTGAAGA 1408
 Db 1499 aaagacagacaagttatga 1518

RESULT 4
 AAT37312
 ID AAT37312 standard; cDNA to mRNA; 1508 BP.

AC AAT37312;

DT 06-FEB-1997 (first entry)

XX Aromatic acyl transferase coding sequence.

DE Aromatic acyl transferase; transformation; anthocyanin pigment;

KW plants; acylation; colour; tone; colouration; colour change;

KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;

KW Senecio cruentus; Lavandula angustifolia; ds.

XX Senecio cruentus (Clone pCAT8).

XX Key Location/Qualifiers

FT 3..1367

FT /*tag= a

FT /product= Aromatic acyl transferase.

XX W09625500-A1.

XX 22-AUG-1996.

XX 16-FEB-1996; 96WO-JP00348.

XX 30-JAN-1996; 96JP-0046534.

PR 17-FEB-1995; 95JP-0067159.

PR 29-JUN-1995; 95JP-0196915.

XX (SUNR) SUNTORY LTD.

XX Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;

PI Nakao M, Tanaka Y, Yonekura K;

XX WPI; 1996-393401/39.

DR P-PSDB; AAW04726.

XX DNA coding for aromatic acyl transferase - for transforming plants

PT which produce anthocyanin pigments and thus altering colour tone,

PT e.g. of flowers

XX Claim 4; Page 69-72; 94pp; Japanese.

XX Vectors containing DNA fragments encoding proteins of plant origin

CC with aromatic acyl transferase activity may be used to transform

CC plants which produce anthocyanin pigments. The aromatic acyl

CC transferase acylates the pigments in the flower resulting in colour

CC tone changes and allowing new colourations to be produced. Six

CC specific DNA sequences encoding aromatic acyl transferase from

CC different plants are described in AAT37308-T37313.

XX

SQ Sequence 1508 BP; 442 A; 294 C; 295 G; 477 T; 0 other;

Query Match 12.9%; Score 190.6; DB 17; Length 1508;
 Best Local Similarity 49.6%; Pred. No. 3.9e-40;
 Matches 676; Conservative 0; Mismatches 654; Indels 33; Gaps 6;

QY 2 CGTGTATCGAAACGCTGTAGAGTTGGGCGCGCGGACTCGGTGGCGAGCAATCGGTGCC 61
 Db 5 catctcgaaacatgccgaataatcgcccccctcgggcaccatcgccatcgctcgttatc 64
 QY 62 GGTACATCTTCGACATGACGTGGCTGCATTTTCATCCATGCTTCAGCTCCTCTCTCTA 121
 Db 65 tcttactttctcgacatttctgctactcttccctccggtccaccatctttcttcttcla 124
 QY 122 CGAATTCCTCTTCCCAAGCAACATTTTCAGATCCATCGTTCCTCAAACTCAAAACAATC 181
 Db 125 tgactttccacattcttaaaatcccatcttcacacatactgttcccgagtaaaacatac 184
 QY 182 TCTCTCTAAACTCTCATACACTTCTTCCCTCTCTCATCAATTTTAATCTACCTTTTCATC 241
 Db 185 tttatcggtcactcttcaacatttttcccggttctgtagtaattgattgatttctctaa 244
 QY 242 CCCGG-----AGAAATGCCGAGTTTCGGTATCTATCTATCCGGGACACTCGGT 286
 Db 245 cactgatggttcgggttttataaaaaaacagaaaataaacacgtttgaagggtgattctgt 304
 QY 287 TTCTTTTCCACTCGCAGAAATCTAGCGACACTTCGATGATCTCGTCGGAATTCGTCAGA 346
 Db 305 tbtggttactttgcagaatgttcttgactttaataattgacagggaaaatcactcctcg 364
 QY 347 ATCTCCGCTTAGGCTCTACAACTTTTGTCCTTAAATTCGCCGCCCATTTGTGCAAGAATCCGA 406
 Db 365 aaaaatgaaaaacttttaccactgtaccttcatttggaatgcaatcaaatattatgtga 424
 QY 407 TAGAAACTCTTCCAAAGTTTTCGCCGTGACGTGACTCTTTTCCAGCGCGAGCGCTCGG 466
 Db 425 ttgcgtcacggtcccaactttttcacttcaagtgaagctttttccgggctcggtatatac 484
 QY 467 TATTGGAATAGCAACGCATCACACGTTAGCGACGCCCTCGTCTCTCGCTTTTATAC 526
 Db 485 actaggaatgacgaatacatcatgacctgtgacgctagcagcggttcaacttttctgaa 544
 QY 527 GGCTTGGTCTTCAATGAGCAACACATTTGAAATGAAGATGAAGATGAAGAATTTAAATC 586
 Db 545 aggggtgacttcgattattcaatctggtgtagatcggtctcttttaacgaagg---atc 601
 QY 587 TTTGCCAGTTTTCGATGATGCTGCTATATAAATATCCGACGAATTTTGACTTCCATTTATG 646
 Db 602 tccaccggtttttgatagat---tgattaaacatccccacatttagatgaataaagttag 658
 QY 647 GAGAAACGCGCTAAATTTTCTTTGCAATCTCGTCACTCTCATACCGACGCGACCGCAT 706
 Db 659 acatacaaggctcgaaaagtcttataaaccttcgagccttctgttggctccactgataaagt 718
 QY 707 TCGAACACAGTTTCGTTTTCACCAATCCAAATTAAGAAATTTGAAGGTTTGGATTTCAGTC 766
 Db 719 tcggtcaacggttctgtagacccgaactaataatcaatctactaaaagaaaaggtcttaac 778
 QY 767 CAGAGTTCCAAAGTTTATGCTCTCTCATCTTTTGTAGCGATTTCAGCTTATATGTGGGC 826
 Db 779 ccaagtgcacaaacttgagtagatgctcttttactgtaacttctgtgttatataatggag 838
 QY 827 TGGCATAACGAATCATTC-----ACACGATGAGAACCAACACACAGCATGCATTCATT 880
 Db 839 ttgcatagcgaataactcactgtaaaataaggaagaaagggcggaagacgagttagaaca 898
 QY 881 TTTCTTGATTCCGGTTCGATCTAAGGCCACGATTAGATCCGCCGGTTCCTGAAATTTACTT 940

Db 855 atgacgtggtacacgtatgacatcgtcgtcaaatcaaaagatgacgtcgtacagag 914
QY 859 -----ACCAAGACACAGGAGTATTTTCTTGATTCGGTCCGATCTAAGGCCA 908
Db 915 gaatcatcgaacgaagaaatagctcagtagtactcagttttacagcggattgcgcagga 974
QY 909 CGATTAGATCCCGCGTTCCTGAAAATTAATTCGGGAACGTGCTTATCGTACGCGCTGC 968
Db 975 cttctgacgccccgtgcgcctaactcttgcaactgtctcgtcgtcgttgca 1034
QY 969 AGAATCCGCGCGGAGAGTGTGGGAGAGAAGGGGTGTTCTGCGCAGCTGAGGTAATC 1028
Db 1035 aaagcaacacataaagagtagtgggataaagggtctcttctgtcagttgcagctatt 1094
QY 1029 GCGCGCGAGATAAAAAAGGAT---CAAGCACAAGAGATATTAGAAACGGTGGAGAA 1085
Db 1095 ggagaagccattgaaaagggttgcaaacgaaaaaggcgtctctgcagatgcaaaaact 1154
QY 1086 TGGTCCCGGAGATTTCGTAAGCGTTGCAGAAATCATATTTTCGGTGGCAGGATCGAGC 1145
Db 1155 tggttatcgaatcctaataatggaatcccttcaaaaagattctcgggattaccggatcgct 1214
QY 1146 AAGCTAGATCTTTACGGTCAGATTTTGGATGGGAGCGGAGAGCAAGCAAAATATTG 1205
Db 1215 aagttcgtatcgtatgtagatttggatgggaaagcctgcgcaaaatttgacattacc 1274
QY 1206 TCGATTGATGGGAGAAATATCGATGACGCTTTGTAAGCCAGGAGATTTCGAGGAGGA 1265
Db 1275 tctgttgattatgcagaattg---atttatgtatcagtcacaggttttgaagaggt 1331
QY 1266 TTGGAGGTTTGTGCTTGTGCTAAGGACAAAATGATGCTTTTCTGCTTATTTTTCA 1325
Db 1332 gtggagattggagatcatctgcctaagatcatatgagatgcatgcttgcacaaatccttga 1391
QY 1326 CTGGGAATTAATGGTTAATAATGATGATTAATTAATAATAATTAATTAATAATAA 1385
Db 1392 gaaggcttcttctcttgcatagtctcttataagaacacatttgcgcaataaagta 1451
QY 1386 TTAAGTGTTCAGTAACTGAGCAATTAATCCCTATTAT 1422
Db 1452 ccaagtccttagtaacacatcaccaaacccacttt 1488

RESULT 6
AAF58252/c
ID AAF58252 standard; DNA; 936 BP.
XX AC AAF58252;
XX DT 24-APR-2001 (first entry)
XX DE Oligonucleotide D1835.
XX EW Electron-transfer group; ETM; mismatch; genotyping;
XX KW gene expression; ss.
XX OS Synthetic.

XX WO200107665-A2.
XX PD 01-FEB-2001.
XX PF 26-JUL-2000; 2000WO-US20476.
XX PR 26-JUL-1999; 99US-0145695.
XX PR 17-MAR-2000; 2000US-0190259.
XX PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX PI Umek RM;
XX WPI; 2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX Example 6; Page 127; 153pp; English.
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;
SQ

Query Match 6.4%; Score 94.8; DB 22; Length 936;
Best Local Similarity 7.1%; Pred. No. 3.5e-15;
Matches 36; Conservative 284; Mismatches 186; Indels 0; Gaps 0;
QY 974 GCGCGCGCAGAGCTGCTGGGAGAGAAAGGGGTGTTCTGCGCAGCTGAGTAATCGCGC 1033
Db 850 GGG 791
QY 1034 GGAGATAAAAAAAGGATCAACGACAGAGAGATATTAGAAACGGTGGAGAAATGGTCGCC 1093
Db 790 GGCWNN 731
QY 1094 GGAGATTCGTAAGCGCTTCAGAAATCATATTTTCGGTGGCAGGATCGAGCAAGCTAGA 1153
Db 730 WNN 671
QY 1154 TCTTTACGCTGCAGATTTTSGATGGGGGAGCGGAGCAAAAGCAAGAAATATTGTCGATTGA 1213
Db 670 WNN 611
QY 1214 TGGGGAGAAATATGCAATGACGCTTTGTAAGCCAGGATTTGCAAGGAGGATTCGAGGT 1273
Db 610 WNN 551
QY 1274 TTGCTTGTCTTTGCCCTAAGCACAATAATGATGCTTTTTCGCTTATTTTTCACCTGGGAAT 1333
Db 550 WNN 491
QY 1334 TAATGCTTAATAATGATGTAATAACTAATAATATTATGTAACAATAATAATTAAGTGT 1393
Db 490 WNN 431
QY 1394 TGAGTAACGTGAAGAATAATCCCTATTATATATATTATGTTTCAATAAAGTGTA 1453
Db 430 WNN 371
QY 1454 AGCTCTTGAAAAAATAAAAAA 1479
Db 370 WNN 345

RESULT 7
AAF58254/c
ID AAF58254 standard; DNA; 936 BP.
XX AC AAF58254;
XX DT 24-APR-2001 (first entry)
XX DE Oligonucleotide D1875.
XX KW Electron-transfer group; ETM; mismatch; genotyping;
XX KW gene expression; ss.
XX OS Synthetic.

Db 730 WWWWWW...
Qy 1154 TCCTTACGGTCGAGATTTTGATGGGGGAAGCGGAGAAGCAAAATAATTGTGCATTGA 1213
Db 670 WWWWWW...
Qy 1214 TGGGGAGAAATATGCAATCACCGTTTGTAAAGCCAGGAGATTTCGAAGGAGGATTTGGAGGT 1273
Db 610 WWWWWW...
Qy 1274 TTGCTTGCTTTGCCCTAACGACAATAATGCATCCTTTTGCCTGCTTATTTTTCACCTGGGAAT 1333
Db 550 WWWWWW...
Qy 1334 TAATGGTTAATAATCATGTATTAATAACTAATAATATTATGTAAACAATTAATTAAGTGT 1393
Db 490 WWWWWW...
Qy 1394 TGAGTAACGTGAAGAATAATCCCTATTATATATTATGATTTGGTTCAAATAAAGTGTAA 1453
Db 430 WWWWWW...
Qy 1454 AGCCTCTTGAAAAAAAAAAAAAA 1479
Db 370 WWWWWW...

RESULT 10
AAF58262/c
ID AAF58262 standard; DNA; 936 BP.
XX AC AAF58262;
XX DT 24-APR-2001 (first entry)
XX DE Oligonucleotide D2007.
XX KW Electron-transfer group; ETM; mismatch; genotyping;
XX KN gene expression; ss.
XX OS Synthetic.
XX PN WO200107665-A2.
XX PD 01-FEB-2001.
XX PF 26-JUL-2000; 2000WO-US20476.
XX PR 26-JUL-1999; 99US-0145695.
XX PR 17-MAR-2000; 2000US-0190259.
XX PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX PI Umek RM;
XX WPI; 2001-159728/16.
XX Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface -
XX Example 6; Page 128; 159pp; English.
XX The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETW) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) CC and single-nucleotide polymorphisms, e.g. for genotyping,
XX monitoring gene expression.
SQ Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other;


```

XX PF 26-JUL-2000; 2000WO-US20476.
XX PR 26-JUL-1999; 99US-0145695.
XX PR 17-MAR-2000; 2000US-0190259.
XX PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX PI Umek RM;
XX DR WPI; 2001-159728/16.
XX XX
XX PT Nucleic acids containing electron-transfer group, useful as labels in
XX PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
XX PT a single surface -
XX PS
XX PS Example 6; Page 127; 159pp; English.
XX CC The present invention relates to a composition comprising two nucleic
XX CC acids each containing an electron-transfer group (ETM) having
XX CC different redox potentials. The invention is used for electronic
XX CC detection of nucleic acids, especially of substitutions (mismatches)
XX CC and single-nucleotide polymorphisms, e.g. for genotyping,
XX CC monitoring gene expression.
XX SQ Sequence 936 BP; A; 139 C; 10 G; 7 T; 776 other:

Query Match          6 28; Score 91; DB 22; Length 936;
Best Local Similarity 0.5%; Pred. No. 3.4e-14;
Matches 4; Conservative 450; Mismatches 305; Indels 0; Gaps

Qy 721 TTTCACCCCAATCAAATAAAGAAATTGAAGGGTTGGATTCCAGTCCAGAGTTCCAAGTT 780
   : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: : :: :
Db 21 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 80

Qy 781 TAGTCCATCTCTCATCTTTTGTAGCGATTGCAGCTTATATGCGCTGGCATAACGAAAT 840
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 81 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 140

Qy 841 CATTCACAGCAGATGAAGACCACAGCACAGGAGTGATTTCTTGATTCGGTTCGATC 900
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 200

Qy 901 TAAGGCCACGATTAGATCCCGCGTTCTGTAATAATTACTTCGGGAACCTGCTATCGTAGC 960
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 201 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 260

Qy 961 CGCTGCCGAGAATCGCGGCGAGAGCTGTGGGAGAGAAAGGGGTGTTCTGGCAGCTG 1020
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 261 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 320

Qy 1021 AGGTAATCGCGCGGAGATAAAAAAAGGATCAACGACAAGAGATATATTAGAACGGTGG 1080
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 321 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 380

Qy 1081 AGAAATGGTCCCGGAGATTCGTAAAGCGTTGCAGAAATCATATTTTCGGTGGCAGGAT 1140
   : : : : : : | : : : : : : : : : : : : : : : : : : : : : : : : :
Db 381 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 440

Qy 1141 CGAGCAAGCTAGATCTTTACCGTGCAGATTTTGGATGGGGAAGCGAGAAAGCAAGAAA 1200
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 441 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 500

Qy 1201 TATTGTCGATTGATGGGGAGAGAAATATGCAATGACGCTTTGTAAGCCAGGGATTTGCAAG 1260
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 501 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 560

Qy 1261 GAGGATTCGAGGTTGCTTGCTTTGCCCTAAGGCACAAAATGGATGCTTTTCTGCTTATT 1320
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 561 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 620

Qy 1321 TTTCACCTGGGAATTAATGGTTTAATAAATGTATGTAATTAACCTAATATTATATGTAACA 1380

```

[illegible]

RESULT 14
AAF58257
ID AAF58257 standard; DNA: 936 BP.

DT 24-APR-2001 (first entry)

XX
DE
Oligonucleotide D1954.

Electron-transfer group; ETM; mismatch; genotyping;
gene expression; ss.

OS Synthetic.

XX
PN
WO200107665-A2.

01-FEB-2001.
PD
AA

26-JUL-2000; 2000WO-US20476.

XX 26-JUL-1999; 99US-0145695.

PR 17-MAR-2000; 2000US-0190259.
XX

XX PA (CLIN-) CLINICAL MICRO SENSORS INC.

PI Umek RM:

XX DR WPI; 2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in PT

PT hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface

PT

XX

PS Example 6; Page 127; 159pp; English.

The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.

Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

Query Match	6.2%	Score 91;	DB 22;	Length 936;
Best Local Similarity	0.5%	Pred. NO. 3.4e-14;		
Matches	4;	Conservative 450;	Mismatches 305;	Indels 0;
Gaps	0;			

QY	721	TTTTTACCCAAATCAAAAATTAAGAAATTAAGAGGTTGGATTTCAGTCCAGAGTTCCAAGTT	780
DB	21	#####	80
QY	781	TAGTCCATCTCCTATCTTTTGTAGCGATTGCAGCTTATATGTGGCGTGGCATACGAAT	840
DB	81	#####	140
QY	841	CATTACAGCAGATGAAGACCAAGACAAGAGGATGCATTTTCTTGATTCGGTTCGATC	900
DB	141	#####	200
QY	901	TAAGGCCAGGATTAGATCCGCGGTCTCTGAAAATTACTTCGGGAACCTGTTATCGTACG	960
DB	201	#####	260
QY	961	CGTGGGAGATGCGCGCGCAGAGCTGTGTGGGAGAGAAAGGGGTGTTCTCGCACGTG	1020
DB	261	#####	320
QY	1021	AGGTAATCGCGGGAGATAAAAAAGGATCAACGACAAGAGATATTAGAAACGGTGG	1080
DB	321	#####	380
QY	1081	AGAAATGGTCGCGGAGATTCGTAAGCGGTTGCAGAAATCATATTTTCGGTGGCAGGAT	1140
DB	381	#####	440
QY	1141	CGACCAAGCTAGATCTTTACGTTGCAGATTTTGGATGGGGAGCGCAGAGCAAGCAAGAA	1200
DB	441	#####	500
QY	1201	TATTGTGATTTGATGGGGAGAAATATGCAATGACGCTTTGTAAAGCCAGGGATTTCGAAG	1260
DB	501	#####	560
QY	1261	GAGGATTGGAGGTTGCTTGTCTTTGCCCTAAGGACAAAATGGATGCTTTTCGCTGCTATT	1320
DB	561	#####	620
QY	1321	TTTCACTGGGAATTAATGGTTAATAAATGTATGTAATTTAAACTAATATTATTATGTAACA	1380
DB	621	#####	680
QY	1381	ATTAAATTAAGTGTGTAGTAACCTGAAGAAATAATCCCTATTATATATTATGATTGCTTC	1440
DB	681	#####	740
QY	1441	AAATAAGTGTAAAGCCCTCTGAAAAAATAAAAAA	1479
DB	741	#####	779

RESULT 15

AAF58259
ID AAF58259 standard; DNA; 936 BP.
AC AAF58259;
XX
24-APR-2001 (first entry)
DT
XX
XX
DE Oligonucleotide D2004.
XX
XX
KW Electron-transfer group; ETW; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN W0200107665-A2.
XX
XX
PD 01-FEB-2001.
XX
XX
PF 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
PA
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
XX Example 6; Page 128; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
XX Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;
SQ

Search completed: November 5, 2001, 18:11:47
Job time: 15442 sec

Query Match 6.2%; Score 91; DB 22; Length 936;
Best Local Similarity 0.5%; Pred. No. 3.4e-14;
Matches 4; Conservative 450; Mismatches 305; Indels 0; Gaps 0;

QY 721 TTTTACCCCAATCCAAATTAAGAAATTGAAGGTTGGATTCCAGTCCAGAGTTCCAGTT 780
DB 21 TTTTACCCCAATCCAAATTAAGAAATTGAAGGTTGGATTCCAGTCCAGAGTTCCAGTT 80
QY 781 TAGTCCATCTCTCATCTTTTGTAGCGATTGCAGCTTATATGTGGCTGGCATAACGAAAT 840
DB 81 TTTTACCCCAATCCAAATTAAGAAATTGAAGGTTGGATTCCAGTCCAGAGTTCCAGTT 140
QY 841 CATTCACAGCAGATGAAGACCAAGACAAACGAGGATGCAATTTTCTTCATTCCTCGATC 900
DB 141 TTTTACCCCAATCCAAATTAAGAAATTGAAGGTTGGATTCCAGTCCAGAGTTCCAGTT 200
QY 901 TAAGGCCACCATAGATCCCGCTTCCCTGAAATTAATCTTCGGGAATGCTTATCGTACG 960
DB 201 TTTTACCCCAATCCAAATTAAGAAATTGAAGGTTGGATTCCAGTCCAGAGTTCCAGTT 260
QY 961 CGTCCGAGAAATCGCGCGCAGAGCTGTGGGAGAGAAAGGGTGTTCCTGGCAGCTG 1020
DB 261 TTTTACCCCAATCCAAATTAAGAAATTGAAGGTTGGATTCCAGTCCAGAGTTCCAGTT 320
QY 1021 AGGTAATCGCGCGGAGATAAAAAAGGATCAACGACACAGAAATATTAGAAACGGTGG 1080
DB 321 TTTTACCCCAATCCAAATTAAGAAATTGAAGGTTGGATTCCAGTCCAGAGTTCCAGTT 380


```
Query Match 4.4%; Score 65.8; DB 1; Length 7218;
Best Local Similarity 7.5%; Pred. No. 9.7e-10;
Matches 31; Conservative 219; Mismatches 161; Indels 0; Gaps 0;

Qy 862 AAGACACGAGGATGCAATTTCTTGATTCGGTGCATCTAAGCCACGATTAGATCGG 921
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1505 AAAACGGCATGTAGCATCACTGTAATACCTATCTATGCAAGTAGTAAAGAGATAGAA 1446

Qy 922 CGGTCTCTGAAATTTACTTCGGGAACGTCTATCTAGTCGGCTCCGAGATCGCGCGC 981
  || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1445 GAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1386

Qy 982 GAGAGCTGGTGGGAGAGAGGGGTGTTCTGGCAGCTGAGGTAATCGCGCGGAGATAA 1041
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1385 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1326

Qy 1042 AAAAAAGATCAACGACAGAGATATAGAACGGTGGGAAATGTCGCCGAGATTC 1101
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1325 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1266

Qy 1102 GTAAGCGTTGCAGAAATCATATTTTCGGTGGCAGGATCGAGCAAGCTAGATCTTTACG 1161
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1265 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1206

Qy 1162 GTCAGATTTTGGATGGGAGCGGAGAGAAAGAAATATTTGTCGATTGATGGGAGA 1221
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1205 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1146

Qy 1222 AATATGCAATGACCTTTGTAACCCAGGATTTGGAAGAGAGATTGGAGG 1272
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1145 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1095

RESULT 2
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
; US-08-232-463-14

Query Match 3.9%; Score 57.2; DB 1; Length 7218;
Best Local Similarity 3.8%; Pred. No. 4.2e-07;
Matches 14; Conservative 213; Mismatches 141; Indels 0; Gaps 0;

Qy 60 CCGCTCACATCTTCGACATGACGTGCTGCATTTTCATCCATGCTCAGCTCCTCTTC 119
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1089 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1148

Qy 120 TAGCAATTCCTTGTTCGAAGCAACATTTTTCAGAATCCATGTTCCAAAACCTCAAA 179
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1149 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1208

Qy 180 TCTCTCTCTAAACTCTCATACATCTTCTCCCTCTCTCATGCAATTTAACTACCTTCA 239
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1209 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1268

Qy 240 TCCCGGAGAAATCGCGAGATTCGCTATCTATCTATCCGGGACTCGGTTCTTTCACCATC 299
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1269 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1328

Qy 300 CGAGAATCTAGCGACGACATCTCGATGATCTGTCGGAATCTCCAGAATCTCCCGTAGG 359
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1329 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1388

Qy 360 CTCTACAACTTTGTCCCTAAATTCGCCGCCCATCTGCAAGAATCCGATAGAAAACCTCTC 419
  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1389 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1448

Qy 420 CAAGTTT 427
  - - - - -
Db 1449 TATCTCT 1456

RESULT 3
US-07-593-657-14
; Sequence 14, Application US/07593657
; Patent No. 5268317
; GENERAL INFORMATION:
; APPLICANT: Tomalski, Michael D.
; APPLICANT: Miller, Lois K.
; TITLE OF INVENTION: INSECT-SPECIFIC PARALYTIC NEUROTOXIN
; TITLE OF INVENTION: GENES FOR USE IN BIOLOGICAL INSECT CONTROL: METHODS AND
; TITLE OF INVENTION: COMPOSITIONS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/593,657
; FILING DATE: 19901004
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
```


; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE: PAG1692RP
; ORGANISM: PAG1692RP
US-08-998-416-1137

Query Match 2.8%; Score 41.6; DB 4; Length 636;
Best Local Similarity 51.6%; Pred. No. 0.0051;
Matches 95; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
Qy 1296 AAAATGGATGCTTTTGGCTTATTTTTCACCTGGGAATTAATGTTAATAATGATGTA 1355
Db 7 AATATATAATTAATTAATTAAGATTAAATAAACTTTTTTATTAATAATTTAAGTA 66
Qy 1356 ATTAACCTAAATATTATTATTAACAAATTAATTAAGTGTGAGTAACGTGAAGAATAATCC 1415
Db 67 TTAATATTAACTATATATATCATTTATTAATAATTAATTTGATTTAATTAATAC 126
Qy 1416 CTATTATATTTATGATTTGGTTCACAAATAAGTGAAGCTCTTGAAAAAATAATCC 1475
Db 127 TTTATATAATATTATATAATTTACTTAAATTCATCATTTATTAATATTTATTAATAT 186
Qy 1476 AAAA 1479
Db 187 AAAA 190

RESULT 6
US-08-998-416-288
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38, 241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CCC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 288:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 837 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE: PAG1241RP
; ORGANISM: PAG1241RP
US-08-998-416-288

Query Match 2.8%; Score 41.6; DB 4; Length 837;
Best Local Similarity 51.6%; Pred. No. 0.0061;
Matches 95; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
Qy 1296 AAAATGGATGCTTTTGGCTTATTTTTCACCTGGGAATTAATGTTAATAATGATGTA 1355
Db 7 AATATATAATTAATTAATTAAGATTAAATAAACTTTTTTATTAATAATTTAAGTA 66
Qy 1356 ATTAACCTAAATATTATTATTAACAAATTAATTAAGTGTGAGTAACGTGAAGAATAATCC 1415
Db 67 TTAATATTAACTATATATATCATTTATTAATAATTAATTTGATTTAATTAATAC 126
Qy 1416 CTATTATATTTATGATTTGGTTCACAAATAAGTGAAGCTCTTGAAAAAATAATCC 1475
Db 127 TTTATATAATATTATATAATTTACTTAAATTCATCATTTATTAATATTTATTAATAT 186
Qy 1476 AAAA 1479
Db 187 AAAA 190

RESULT 7
US-07-593-657-6
; Sequence 6, Application US/07593657
; Patent No. 5266317
; GENERAL INFORMATION:
; APPLICANT: Tomalski, Michael D.
; APPLICANT: Miller, Lois K.
; TITLE OF INVENTION: INSECT-SPECIFIC PARALYTIC NEUROTOXIN
; TITLE OF INVENTION: GENES FOR USE IN BIOLOGICAL INSECT CONTROL: METHODS AND
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C. 201
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/593,657
; FILING DATE: 19901004
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33, 878
; REFERENCE/DOCKET NUMBER: 14-90
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1241 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double


```

RESULT 11
US-08-319-704-5
; Sequence 5, Application US/08319704
; Patent No. 5814617
; GENERAL INFORMATION:
; APPLICANT: Hoffman, Stephen L.
; APPLICANT: Charoenvit, Yupin
; APPLICANT: Charoenvit, Richard C.
; APPLICANT: Hedstrom, Richard C.
; APPLICANT: Doolan, Denise L.
; TITLE OF INVENTION: Protective 17 kDa Malaria Hepatic and
; TITLE OF INVENTION: Erythrocytic Stage Immunogen and Gene
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:

```

```

RESULT 12
US-08-415-593-40
; Sequence 40, Application US/08415593
; Patent No. 5912140
; Patent No. 5912140 5776726
; GENERAL INFORMATION:
; APPLICANT: Whoriskey, Susan K.
; APPLICANT: Quinn, Cheryl L.
; APPLICANT: Tao, Nisajun
; APPLICANT: Politis-Virk, Karen I.
; APPLICANT: Schimmel, Paul R.
; TITLE OF INVENTION: Recombinant Pneumocystis Aminoacyl tRNA
; TITLE OF INVENTION: Synthetase Genes, Tester Strains and Assays
; NUMBER OF SEQUENCES: 46
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: U.S.

```

ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/415,593
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI94-09
TELEPHONE: 617-861-6240
TELEFAX: 617-862-9540
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 1991 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 29..1780
US-08-415-593-40

Query Match 2.5%; Score 36.4; DB 2; Length 1991;
Best Local Similarity 48.5%; Pred. No. 0.43;
Matches 100; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 1274 TTGCTTGTCTTGGCCCTAAGCAGCAAAATGATGCTTTTGGCTGCTTATTTTTCACCTGGGAAT 1333
Db 1780 TTGATTAATTTTAAATAAATAATTTTGAATAATAATTTTTCAAAATAATACTCTTATT 1839
QY 1334 TAATGCTTAATAATCTATGTAATAAATAATAATTAATAATAATAATAATAATAAGTGT 1393
Db 1840 TTGTCATATATTTTGAATAAATAATAATAATAATAATAATAATAATAATAATAATAAT 1899
QY 1394 TGAGTACGCTGAAGATAATCCCTATATATATTTATGATTTTGGTTTCAATAAAGTGTA 1453
Db 1900 ACTTAAATATATTAAGTGATTTTAAAGCTTATAAATTTCTTGTAAAATAATCTATG 1959
QY 1454 AGCCTCTTGAAGAAAAAATAAAAAA 1479
Db 1960 TATGTATTAAAAAATAAAAAAATAAAAAA 1985

RESULT 13
US-08-415-593-39
Sequence 39, Application US/08/415593
Patent No. 5912140
Patent No. 5912140 5776726
GENERAL INFORMATION:
APPLICANT: Whoriskey, Susan K.
APPLICANT: Quinn, Cheryl L.
APPLICANT: Tao, Nijun
APPLICANT: Politis-Virk, Karen I.
APPLICANT: Schimmel, Paul R.
TITLE OF INVENTION: Recombinant Pneumocystis Aminoacyl tRNA
TITLE OF INVENTION: Synthetase Genes, Tester Strains and Assays
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.
ZIP: 02173
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/415,593
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI94-09
TELEPHONE: 617-861-6240
TELEFAX: 617-862-9540
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 2128 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
US-08-415-593-39
Query Match 2.5%; Score 36.4; DB 2; Length 2128;
Best Local Similarity 48.5%; Pred. No. 0.45;
Matches 100; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 1274 TTGCTTGTCTTGGCCCTAAGCAGCAAAATGATGCTTTTGGCTGCTTATTTTTCACCTGGGAAT 1333
Db 1910 TTGATTAATTTTAAATAAATAATTTTGAATAATAATTTTTCAAAATAATACTCTTATT 1969
QY 1334 TAATGCTTAATAATCTATGTAATAAATAATAATTAATAATAATAATAATAATAAGTGT 1393
Db 1970 TTGTCATATATTTTGAATAAATAATAATAATAATAATAATAATAATAATAATAATAAT 2029
QY 1394 TGAGTACGCTGAAGATAATCCCTATATATATTTATGATTTTGGTTTCAATAAAGTGTA 1453
Db 2030 ACTTAAATATATTAAGTGATTTTAAAGCTTATAAATTTCTTGTAAAATAATCTATG 2089
QY 1454 AGCCTCTTGAAGAAAAAATAAAAAA 1479
Db 2090 TATGTATTAAAAAATAAAAAAATAAAAAA 2115

RESULT 14
US-08-378-235B-8/c
Sequence 8, Application US/08378235B
Patent No. 5872241
GENERAL INFORMATION:
APPLICANT: Pyle, Anna M.
APPLICANT: Michels, William J.
TITLE OF INVENTION: MULTIPLE COMPONENT RNA CATALYSTS AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper and Dunham, LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10032
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/378,235B
FILING DATE: 25-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/45553
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA (genomic)
US-08-378-2358-8

Query Match 2.4%; Score 36; DB 2; Length 174;
Best Local Similarity 58.3%; Pred. No. 0.11;
Matches 63; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
Qy 1343 ATAAATGATCTAATTAACCTAATATTATTATGTAACAATTAATTAAGTGTGAGTAACG 1402
Db 136 ATATGTCGAATCAATTTTCGTTATTATTATTATTAAATTAATTAATTAATTAATA 77
Qy 1403 TGAAGAATAATCCCTATTATATATTTATGATTGTTGTTCAATAAAGTG 1450
Db 76 ATATTAACCATTTATTATATATTTTTTTATTAATAATATATAATATTG 29

RESULT 15
US-07-867-106-2
Sequence 2, Application US/07867106
Patent No. 5389526
GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 19920625
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
APPLICATION NUMBER: PCT/AU90/00530
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Feeney, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
US-07-867-106-2

Query Match 2.4%; Score 36; DB 1; Length 5852;
Best Local Similarity 47.4%; Pred. No. 1.2;
Matches 108; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
Qy 1252 ATTTTCGAGGAGGATTGGAGGTTTCTGCTTTTCCCTAAGACAAAATGGATGCTTTTG 1311
Db 1814 ATTTTAAATAAATCTACTTTTTTTTTTTTTTTTTTTTTTTTAAATTTTAAATTTT 1873
Qy 1312 CTGCTTATTTTTCACCTGGGAATTAATGCTTAATAAATGATGTAATTAACCTAATATAT 1371
Db 1874 TTTTATTAGATCTCATAATTAATAATCAATTAATAATTAAGTATTTTAAATAT 1933
Qy 1372 TATGTACAATAATTAAGTCTTGAGTAACGTGAAGATAATCCCTATATATATTTATG 1431
Db 1934 GCAAAAACATATAAAAACTAATAGTTTAAACCAACTTTTTTCTATTTCTTTTTTTT 1993
Qy 1432 ATTTGGTTCAATAAAGTGAAGCCCTCTTGAAAAAATAAAAAA 1479
Db 1994 TTTTTTTTTTTACTTTGAAAAAATAAAAAAATAAAAAA 2041

Search completed: November 5, 2001, 18:04:45
Job time: 15085 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:00:51 ; Search time 5816.79 Seconds
(without alignments)
2403.520 Million cell updates/sec

Title: US-08-894-356C-4
Perfect score: 1479
Sequence: 1 CCGTGATCGAAACGCTGTAGA.....TTGAAAAAAAAAAAAAAAAAAAA 1479

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	139.4	9.4	687	111	AW102336	s886d06.y
2	131.8	8.9	690	155	BG599447	EST504342
3	131.4	8.9	606	113	AW221049	EST297518
4	130.8	8.8	783	155	BG584708	EST486469
5	129.6	8.8	570	142	BE922784	EST426553
6	127.8	8.6	583	113	AW221050	EST297519
7	127.8	8.6	763	155	BG591758	EST499600
8	118.2	8.0	508	173	BG097054	EST461573
9	112.4	7.6	564	142	BE921494	EST425179
10	112.4	7.6	559	155	BG598154	EST496832
11	109.8	7.4	788	164	BE187619	EST336180
12	108	7.3	574	142	BE922746	EST426515
13	106.6	7.2	482	144	BF096979	EST400516
14	106.6	7.2	591	167	BE434257	EST405335
15	104.8	7.1	487	118	AW616206	EST307245
16	104.4	7.1	575	142	BE920468	EST424237
17	94.8	6.4	518	119	AW650280	EST328734
18	94.8	6.4	563	152	BG362734	sac08a08
19	93.2	6.3	603	118	AW616119	EST296884
20	92	6.2	686	120	AW776512	EST335577
21	91.2	6.2	403	143	BF053202	EST438432
22	88.6	6.0	648	134	BG457806	BF034C06P
23	88.4	6.0	708	138	BE660849	7-C8 Gmax
24	87.8	5.9	497	122	AW930934	EST356777
25	85.8	5.8	367	142	BE920010	EST423779
26	85.4	5.8	368	142	BE921953	EST425722
27	85	5.7	411	142	BE922378	EST426147
28	85	5.7	414	142	BE921952	EST425721
29	82.6	5.6	654	151	BF642724	NF0708031
30	82.2	5.6	656	119	AW692527	NF0560055
31	82.2	5.6	677	118	AW586972	EST318595
32	80.8	5.5	597	120	AW774948	EST334099
33	79.6	5.4	630	155	BG581249	EST482982
34	79.4	5.4	386	173	BG097680	EST462199
35	79.4	5.4	422	155	BG589694	EST497536
36	77.4	5.2	456	142	BE923572	EST427341
37	76.6	5.2	490	110	AW038515	EST280198
38	76.2	5.2	694	117	AW560686	EST315734
39	76.2	5.2	829	155	BG584678	EST486439
40	75.2	5.1	409	111	AW093479	EST286659
41	72.6	4.9	404	120	AW737239	EST338666
42	72.6	4.9	543	119	AW650650	EST329104
43	72.6	4.9	560	30	AV442325	AV442325
44	69.8	4.7	403	143	BF052576	BF052576
45	69.8	4.7	649	151	BF643475	BF643475

ALIGNMENTS

RESULT	1
AW102336	
LOCUS	sd86d06.y1
DEFINITION	Gm-c1009 Glycine max CDNA clone GENOME SYSTEMS CLONE ID: Gm-c1009-1284 5' similar to TR:Q92WB4 Q92WB4 F21M11.13 PROTEIN. ; mRNA sequence.
ACCESSION	AW102336
VERSION	AW102336.1
KEYWORDS	EST
SOURCE	soybean.
ORGANISM	Glycine max
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine. 1 (bases 1 to 687)
AUTHORS	Shoenmaker, R., Keim, P., Vodkin, J., Erpelting, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,

TITLE
JOURNAL
COMMENT

Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoenmaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
Insert Length: 2097 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 431.

FEATURES
source

1. .687
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1009-1284"
/clone_lib="Gm-c1009"
/lab_host="XL10-Gold"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; The mRNA was isolated from entire roots of 2-month-old 'Williams' plants that were greenhouse grown in 5-gallon pots. To suppress nodulation, Black Gold All-purpose potting soil was supplemented with: 0.36g/L available phosphoric acid (P2O5), 20mg/L urea N, 0.16g/L S, 0.49mg/L B, 2.5mg/L Cu, 0.15g/L Fe, 13.53mg/L Mn, 0.26mg/L Mo, 14mg/L Zn, 20mg/L Ca, and the following nutrients in a slow-release form (Osmocote): 0.165g/L ammoniacal N, 0.185g/L nitrate N, 0.35g/L available phosphoric acid, and 0.35g/L soluble potash. No nodules were visible on the roots at harvest. Stratagene's CDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA is hemimethylated. Stratagene's first-strand synthesis primer was used [GAGAGAGAGAGAGAGAGACTAGTCTCGAG(T)-18]. After second-strand synthesis is, the cDNA ends were 'polished' with clone's Pfu DNA polymerase, ligated to EcoRI adapters, and phosphorylated. The XhoI site within the first-strand synthesis primer was restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 400bp cutoff, using a SizeSep 400 Spin column from Pharmacia. The column eluent was then ligated into Stratagene's pBluescript II XR predigested vector (pBluescript II SK(+)) that had been digested with EcoRI and XhoI, and phosphorylated. Both the white and blue colonies appear to contain recombinant plasmids with Kaim and Dr. Virginia Coryell."

BASE COUNT 157 a 247 c 91 g 187 t
ORIGIN

Query Match 9.4%; Score 139.4; DB 111; Length 687;
Best Local Similarity 52.6%; Pred. No. 2.4e-25;
Matches 321; Conservative 0; Mismatches 286; Indels 3; Gaps 1;

QY 16 GTAGAGTGGCGCGCGGACTCGTGGCGAGCAATCGGTGGCTCACATTTCTCG 75

DB 77 GTGAAGTTGTTCACACACCA:GCTCACTTCCCTCACACCCCTTCCCTCACATTTCTCG 136

QY 76 ACATGACGTGGTGCATTTTTCATCCCATGCTTCAGCTCCCTCTTCTACGAATTCCTTGT 135

```

Db 137 ATATTCCAGGTGTTACTGCGCACCCCATCCACGCGATCTTCTCTATGACTTTCCTCCACC 196
Qy 136 CCAAGCAACATTTTTCAGAAATCCATGTTCCAAAACCTCAAAACATCTCTCTCTAAACATC 195
Db 137 CCACACCACTTTCTCCAAACAGCACTTCCCATCTCAACACATCTCTTCCCTTAACCC 256
Qy 196 TCATACACTTCTTCCCTCTCTCATGCAATTTAAATTAACCTTCACTCCCGGAGAAATGCG 255
Db 257 TCCAAACACTTCTCCCTCTCTCCAAATCTCATGTTCTCCACAGCAACCCCATCTCT 316
Qy 256 CGGAGTTTCGTAATCTATCCGGGACTCGGTTCTTTCACCATCCGAGAAATCTAGCGAGC 315
Db 317 CTCACATACGCTACCTTGATGAGACTCTCTCTCTCCAGGTTGCAGAGTCCACCGCAG 376
Qy 316 ACTTCGATGATCTGTCGGAATCGTCCAGAAATCTCCCGTTAGGCTCTACAACTTTTGCC 375
Db 377 ACTTCACCTCTCATACATCAGATTCACCAACAGAGAGTTCCAAATGGCACCCTCTTGTC 436
Qy 376 CTAAATTCGCGCCCATTTCTGGAAGATCCGATAGAAAC-----TCTTCCAAATTTTCGCGC 432
Db 437 CCGCGTTTCTTACCCTACCGCTGTTGACCAAGATGGCGCACGTGTCTTCTCTTATGGCCA 496
Qy 433 TGCAGTGACTCTTTTCCAGCGCGAGCGTGGTATTTGGAATAGCAACGATCACACCG 492
Db 497 TTCAGTTACAAATTTTCCCAAGTCTGGCTTCACCATATGCTCACCTTCAACCACTTG 556
Qy 493 TTAGCGACGCGCGTCTCTCGGTTTATAAGCGCTTGGTCTTCAATGAGCAACACA 552
Db 557 CCAGCGAGCGCAAGTCACTTCAACATTCATCAAGTCTGGGCTCTCTTTGCAAGCAN 616
Qy 553 TTGAAATGAAGATGAAGATGAAGATTTAAATCTTTGCCAGTTTTCGATAGATCCGTC 612
Db 617 AAGGAACATGGCTNNCTTCAACCTCTGCTCTCTTCTCTCATGAGAGGGACAAG 676
Qy 613 TAAATATCC 622
Db 677 TTAAGATCC 686

RESULT 2
BG599447
LOCUS EST504342 cSTS Solanum tuberosum cDNA clone cSTS5122 5' sequence, 12-APR-2001
DEFINITION mRNA sequence.
ACCESSION BG599447
VERSION BG599447.1 GI:13616583
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 690)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chienlingo,A.,
Bougri,O., Bueli,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.
FEATURES
Location/Qualifiers
source 1..690
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS5122"
/clone_lib="cSTS"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

```

XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."

BASE COUNT 178 a 156 c 144 g 212 t

Query Match 8.9%; Score 131.8; DB 155; Length 690;
 Best Local Similarity 52.5%; Pred. No. 2.2e-23;
 Matches 289; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

```

Qy 15 TGTAGAGTTGGCGCGCGGACTCGTGGCGGAGCAATCGGTCCCGCTCACATTCCTC 74
Db 6 TGTCAAGTTGCGCCACCTCCGCGCGCGGAGGAGGACACCTCCCTCTTACTTATTTT 65
Qy 75 GACATGAGTGGGTGCAATTTTCATCCCATGCTTACGCTCTCTTCTAGCAATTCCTTGT 134
Db 66 GATCATGTTTGGTTAGGTTCCACCGGTATAGGGGGATATTATTCTACAAGCTCCCAT 125
Qy 135 TCCAAGCAACATTTTTCAGAAATCGTTCCAAACTCAAAACATCTCTCTAAAAC 194
Db 126 TCCAACCCGATTTGTTCAAAACATTTATCTCTCTTAAAAATCACTCTCCCTCACT 185
Qy 195 CTCATACACTTCTTCTCTCTCATGCAATTTAAATCTACCTTTCATCCCGGAGAAAAATG 254
Db 186 CTCAAACACTATACGCCCTTAGCTGGAACGTTGCTCGCTCACTAGATACAACGGATAT 245
Qy 255 CCGAGATTTCCGTTATCTATCCGGGACTCGGTTTCTTTTCCACCATCGCAGAAATCTAGCAG 314
Db 246 CTTGATGTTTATGTTGACAGGAGATTTCTATCTTCTTCTTCTTCTGAGATGATATG 305
Qy 315 GACTTCGATGATCTCGTCGGAAATCGTCCAGAAATCTCCGTTAGGCTCTACAACCTTGT 374
Db 306 GATTTCAATTTATCTCATTTGGTGACCATCCGCGTAATGCTAAGGATTTTATCACTTTGT 365
Qy 375 CCTAAATTCGCGCGCATTTGTCGAAGAATCCGATAGAAACCTCTTCCAAATTTTCCCGCTG 434
Db 366 CTAGTTAGGAGAACCTTAAGATGCAACCGGGTCCAAATAGCTCCCGCTTACCCAT 425
Qy 435 CAGTGACTCTTTTCCCGAGCGGCTCGGTATTTGGAATAGCAACGATCACACCGTT 494
Db 426 CAAGTGACACTTTTCCGAATCTTGGTGTATCCATTTGTTTCTTCTTCTTCTTCTTCTTCT 485
Qy 495 AGCAGACCCCGCTGTTTCTCGGTTTATACGGCTTTCGCTTCAATGAGCAACACAT 554
Db 486 GGTATGAGCTACCATAGTAGGTTTCATTAGGGCGTGGGCTCTACTCCATAAATTCGGT 545
Qy 555 GAAATGAAGA 565
Db 546 GGAGATGAACA 556

RESULT 3
AW221049
LOCUS EST297518 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION cDNA clone cLEF3F1, mRNA sequence.
ACCESSION AW221049
VERSION AW221049.1 GI:6532733
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
1 (bases 1 to 606)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ann,S., Ronning,C.M.,
Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue

```

JOURNAL COMMENT

Unpublished (1999)
Contact: David Frisch
Clemson University
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel.: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.

FEATURES

source
1..606
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF3F1"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2: XhoI; cLEF - Fruit were tagged at the lcm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"
BASE COUNT 147 a 143 c 123 g 193 t
ORIGIN

Query Match 8.9%; Score 131.4; DB 113; Length 606;
Best Local Similarity 52.0%; Pred. No. 2.8e-23;
Matches 294; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 1 CCGTGATCGAAGCTGTAGATTGGCCGCCCGGACCTCGGTGGCGGAGCAATCGGTGC 60
DB 11 CCGTGATTGAGCAATGTCAGTTGGCCGCCACCTCCCGCGCGCGCAACGGAGGTGATACTCC 70
QY 61 CGTCTACATCTTCGACATGACGTGCTGCTATTTTCATCCATCGTTCAGCTCCTCTCT 120
DB 71 CTTCTACTATTTTGACCATGTTGGTTAGGTTTCGCCGTATGAGCGGATATATTTT 130
QY 121 ACGAATTCCTGTTCCAGCAACATTTTTCAGAAATCCATCGTTCAGAACTCAAAACAT 180
DB 131 ACAAGCTCTCCATTTTCAAAACCGATTTGCTTCAAAACATTTATCTCTCTTAAAAAT 190
QY 181 CTCTCTTAAACTCTCATACACTTCTCCCTCTCTCATGCATTTAATCTACCTTCAT 240
DB 191 CACTCTCCCTCACTCTCAAAACACTATACGCCCTTAGCCGGAACGTTGCTTGCCACTAG 250
QY 241 CCGCGGAGAAAATGCGGAGTTTCGGTATCTATCCGGGACTCGGTTCCTTCCACCATCG 300
DB 251 ATACAAACGATATCTGAGTTAGTATGTGACAGAGATTTCTGTCTGTACTTTT 310
QY 301 CAGAAATCTAGCGAGACTTCGATGATCTCGTCGGAATCGTCCAGAAATCTCCGTTAGGC 360
DB 311 TCAGAGCTGATATGAATTAATTAATCTCATTTGTTGATGACCATCGCGCTAAGGCTAAGGAT 370
QY 361 TCTACAACTTTGCTTAAATTCGCCCCCATTTGTCGAAGAAATCCGATAGAAAATCTTCC 420
DB 371 TTATACATTTGTTCTTAAGTTAGGGAACCTAAGGATGACCGGGGGTCCAACTAGGCC 430
QY 421 AAGTTTTTCGCGGTGACGTTCTTTTCCAGCGCGGCGGTGCGGTATTTGGAATAGCAA 480
DB 431 CGCTCTTAGCCATTCAGGTGACACTTTTCCGAATCTTGGTGTATCCATTTGGTTCACTA 490
QY 481 CGCATCACACCGTTAGCGAGCGCCCGTTCGCGTTTATACGGGTTTGGTCTTCAA 540
DB 491 ACCATCATGTTGTTGTTAGTGAGCTACTATAGCAGGGTTTCATTAAAGCGGTGGCTCTAC 550
QY 541 TGAGCAACACATTTGAATGAAGA 565
DB 551 TCCACAAATTCGGTGGACATGAACA 575

RESULT 4

LOCUS BG584708 783 bp mRNA EST 11-APR-2001
DEFINITION EST486469 MHAM Medicago truncatula/Glomus versiforme mixed EST library cDNA clone PMHAM-19K14 5' end, mRNA sequence.
ACCESSION BG584708
VERSION BG584708.1 GI:13599772
KEYWORDS EST;
SOURCE Medicago truncatula/Glomus versiforme mixed EST library.
ORGANISM Medicago truncatula/Glomus versiforme mixed EST library
REFERENCE 1 (bases 1 to 783)
AUTHORS Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J. and Fraser,C.M.
TITLE ESTs from roots of Medicago truncatula after colonization with Glomus versiforme, 2001
JOURNAL Unpublished (2001)
COMMENT Contact: Harrison M.J.
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401
Tel: 580-223-5810
Fax: 580-221-7380
Email: mjharrison@noble.org
Noble EST name: N380179e TIGR sequence name: MTDBT67TK More information is available at: http://www.medicago.org
Seq primer: SKmod (CTA GAA CTA 9tg GAT CC).

FEATURES

source
1..783
Location/Qualifiers
/organism="Medicago truncatula/Glomus versiforme mixed EST library"
/cultivar="Medicago truncatula genotype A17"
/db_xref="taxon:119092"
/clone="PMHAM-19K14"
/clone_lib="MHAM"
/tissue_type="roots colonized with Glomus versiforme"
/dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days post-inoculation with Glomus versiforme. The library was made from a mixture of RNA from each of these stages."
/lab_host="E. coli strain XL0LR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA from roots harvested at 10, 17, 22, 31 and 38 days post-inoculation with Glomus versiforme. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0LR cells."
BASE COUNT 231 a 219 c 97 g 236 t
ORIGIN

Query Match 8.8%; Score 130.8; DB 155; Length 783;
Best Local Similarity 53.3%; Pred. No. 4.2e-23;
Matches 323; Conservative 0; Mismatches 277; Indels 6; Gaps 2;

QY 3 GTGATGGAACGTTGATAGTTGGCCGCCCGGACTCGGTGGCGGAGCAATC---GGTG 59
DB 9 GTGATGGAACATAGTCAAGTTGGCAACCAAAATTCATTACATCAACCAATTTCTA 68
QY 60 CCGCTCACATTTCTCGACATGACGTGGCTGCTATTTTCATCCCATGCTTACGCTCTTC 119
DB 69 CCCTTAATTTCTTCGACATACCATGGTTTATTGTGACCAATTCACGATATCTTCTTC 128
QY 120 TAGCAATTCCTTGTTCGAAGCAACATTTTTCAGAAATCCATCGTTCGAAGCAACAA 179
DB 129 TACCAATTTCTCATCAACTCATCATTTCTTCTCAAAACAACTTCCCATCTTCAACAC 188
QY 180 TCCTCTCTAAACTCTCATACACTTCTCCCTCTCTCATGCAATTTAATCTACCTTCA 239
DB 189 TCACCTTCCATAAACCCCTCCCTACACTTCTCCCTTTTCTTCCCTTAACCTCATCTCCCTCCA 248

```
QY 240 ---TCCCGGAGAAAATGCCGGAGTTTCGGGTATCTATCCGGGACTCGGTTTCCTTTACCC 296
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 249 AATTCCTCCAAAACACCGATCCTTACATAGTTACCTCGATGAAGACTCTATCTCTTCCACC 308

QY 297 ATCGCAGAAATCAGCGAGCACTTCGATGATCTCGCGGAAATCGTCCAGAAATCTCCCGTT 356
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 309 GTAGTGAATCCTCAGCAGACTTCAATATCTCTGATCTGATTCACAAGATGCTCAAAAAC 368

QY 357 AGGCTCTCAAACTTTGTCCTCCCTAAATTCGGCGCCATTTGCGAAGATCCGATAGAAAATC 416
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 369 TGGCACCATCTGTTCCAAATTTACCTCACTCGCTGACCGAACAAGACAACTCGCATC 428

QY 417 TTCCAAGTTTTCGGCGTGCAGTGACTCTTTTCCAGCGCCGAGCGTGGTATTAAGGAATA 476
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 429 ATCCCTATAATGGCTATTCAAGTCACTGCTTTCCAAACTCGGCGTCTCTATTATTCGCTA 488

QY 477 GCAACGCATCACACCGTTAGCGACCGCCGCTGTTTCGCGGTTTATACGGCTTGGTCT 536
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 489 AGTTACAACCATGTTGCTGCGACGCGAAATACATCACTACCATTTTCGTAATTTTGGGCC 548

QY 537 TCAATGAGCAAAACATTTGAATGAAGATGAAGATTTTAAATCTTTGCGCACTT 596
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 549 TCTCTTTCAAAAGCCATAGCAAAAACAAATAATTTGCTCTTTGAATCTCTCCGCCT 608

QY 597 TTCGAT 602
      |||||
Db 609 CTTGAT 614

RESULT 5
BE922784
LOCUS BE922784 570 bp mRNA EST 02-OCT-2000
DEFINITION EST426553 potato leaves and petioles Solanum tuberosum cDNA clone
CSTB22E2 5' sequence, mRNA sequence.
ACCESSION BE922784
VERSION BE922784.1 GI:10448860
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
Generation of ESTs from potato leaves and petioles
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
Location/Qualifiers
1..570
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="CSTB22E2"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site1: EcoRI; Site2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 139 a 136 c 115 g 180 t
ORIGIN
```

```
Query Match 8.8%; Score 129.6; DB 142; Length 570;
Best Local Similarity 52.6%; Pred. No. 8e-23;
Matches 282; Conservative 0; Mismatches 254; Indels 0; Gaps 0;

QY 1 CCGTCATCAAAACGTGTAGAGTTGGCGCCCGCCCGGACTCGTGGCGGAGCAATCGGTGC 60
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 31 CTTGATAGACAATGTCAAGTTGCGCACCTCCCGCGGCGCAGCGAGGTGACACTCC 90

QY 61 CGCTCACATCTTCACATGACGTGGCTGCAATTTTCATCCCATGCTTCAGTCTCCTCTTCT 120
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 91 CTCCTACTATTTTGCATCATGTTGGTTAGGTTCCACGATATAGCGGAGATATTATTCT 150

QY 121 ACGAATTCCTTGTTCGAAGCAACATTTTTCAGAAATCCATCGTTCCAAACATCAAAACAT 180
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 151 ACAAGCTCCCATTTTCCAAACCCGATTTTCGTTCAAAACATATTCTCTCTTAAAAAT 210

QY 181 CTCCTCTAAAACTCTATACACTTCTCTCCCTCTCTCATGCAATTTAATCTACCTTCTAT 240
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 211 CACTCTCCCTCACTCTCAACACTATAGCGCCTTACGTGGAACGTTGCTCGTCCACTAG 270

QY 241 CCCGGAGAAAATCGCGAGTTTCGGTATCTATCCGGGACTCGGTTTCTTTCACCATCG 300
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 271 ATCAAAACGGATATCCTGAGTTACATTTATGTGACAGGAGATTCTGTATCTGTACTTTT 330

QY 301 CAGAACTCTAGGACGACTTCGATGATCTCGTGGGAAATCGTCAGAAATCTCCCGTTAGGC 360
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 331 CTGAGACTGATATGATTTCAATTTATCTATTGTTGACCATCCGCGGTAATCTAAGGATT 390

QY 361 TCTACAACCTTTGTCCTAAATTCGCGCCCATTTGTCGAGAATCCGATAGAAAATCTTCC 420
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 391 TTTATCACTTTGTTCTAGTTAGGGGAACCTTAAGGATGCACCCGGGGTCCAAATAGCTC 450

QY 421 AAGTTTTCGCGTGCAGTGACTCTTTTCCAGGCGGAGGCTCGGTATTGGAATAGCAA 480
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 451 CGGCTTAGCCATTCAGTGACACTTTTTCGAATCTTTGGTGATCCATTTGGTTCACTA 510

QY 481 CGCATCACACCGTTAGCGACCGCCGCTGTTCTCGCGTTTATACGGCTTGGTCT 536
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 511 ACCATCATGTTGTTGTCATGGAGCTACCATAGTAGGTTTCATTAGGGCGTGGGCT 566

RESULT 6
AW221050
LOCUS AW221050 583 bp mRNA EST 07-DEC-1999
DEFINITION EST297519 tomato fruit mature green, TAMU Lycopersicon esculentum
cDNA clone cLEF3f3, mRNA sequence.
ACCESSION AW221050
VERSION AW221050.1 GI:6532734
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.,
Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
Location/Qualifiers
1..583
/organism="Lycopersicon esculentum"
/cultivar="TA496"
```


Db 166 TTGGTTCAAAACCAATATTCTACTCTTAAACATTTCACTCTCCCTCACTCTCAAAACACTAC 225
 Qy 207 TTCCCTCTCTCATGCAATTAATCTACCTTCTATCCCTTCTATCCCGGAGAAATGCGGAGTTCCG 266
 Db 226 ACACCTTTAGCTGGCAACATGCTGTGTCCACATAATTCGAGTGGTATCTCTGAGTTGCGT 285
 Qy 267 TATCTATCCCGGAGCTCGGTTCTTTTCACCATCGCAGAAATCTAGCAGCACTTCGATGAT 326
 Db 286 TATGTGACTGGAGATCTATATCTGTACTTTTATGAGACTGATATGAATTTCAATCAT 345
 Qy 327 CTCGTCGGAATTCGTCGGAATTCCTCCGTTAGCTCTACAACCTTTGTCCTAAATTCGCG 386
 Db 346 CTCATTTGGTAACCATCTCTGAAAT---GCTAAGGATTTTATCCCAATCTCTCAATTTGGCA 402
 Qy 387 CCATTTGTCGAAGATCCGATGAAATCTTCCAGTTTTCGCGTGCAGTGACTCTT 446
 Db 403 CAACCTTAAGATGACACGGGGTCAAACTAGTCCCGGTCTTACCCATTTCAAGTGACACTT 462
 Qy 447 TTCCAGGCGGAGCGCTCGGTATTTGAATAGCAACGCATCACACCGTTAGCGACGCGCG 506
 Db 463 TTCCGAATCTTGGCATATCCGTTGGTTTATGTAACCATCATCTCGCTTGTGATGGAAT 522
 Qy 507 TCGTTTCTCGGTTTATACGCGTTGGTCTTCAATGAGCAAA 548
 Db 523 ACCATCGTGAATTTCAAGAACATGGGTTTACTCAACAAA 564

RESULT 10
 BG598154 659 bp mRNA EST 12-APR-2001
 LOCUS EST496832 cSTS Solanum tuberosum cDNA clone cSTS20M23 5' sequence,
 DEFINITION mRNA sequence.
 ACCESSION BG598154
 VERSION BG598154.1 GI:13616294
 KEYWORDS EST.
 SOURCE potato.
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 659)
 van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chienmingo,A.,
 Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
 Generations of ESTs from sprouting potato eyes
 Unpublished (2000)
 Contact: Cathy Ronning
 The Institute for Genomic Research
 For clone info: please contact Research Genetics, Libraries
 Division tel 1-800-711-6195, email cdna@resgen.com
 Seq primer: M13P-R.
 FEATURES Location/Qualifiers
 source
 1. .659
 /organism="Solanum tuberosum"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cSTS20M23"
 /clone_lib="cSTS"
 /tissue_type="sprouting eyes from tubers"
 /dev_stage="12-14 weeks post harvest"
 /lab_host="SOLR"
 /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
 XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
 taken from tubers. The tubers were incubated at 26C in the
 dark for 2-3 weeks prior to sprouting. The eyes were
 frozen in liquid nitrogen immediately upon removal from
 tubers."

BASE COUNT 178 a 163 c 112 g 206 t
 ORIGIN
 Query Match 7.6%; Score 112.4; DB 155; Length 659;
 Best Local Similarity 53.0%; Pred. No. 2.3e-18;

Matches 287; Conservative 0; Mismatches 251; Indels 4; Gaps 2;
 Qy 27 CCGCGCCGGAGCTCGGTGGCGGAGCAATCGGTGGCGCTCACATTTCTTCGACATGAGCTGG 86
 Db 113 CTCCCGCAGGGCGGCAACAGAGCTAACGCTCCCTCTAATTTATTTGATCT-ATATGG 171
 Qy 87 CTGCAATTTTATCCCATGCTTCAGCTCTCTTCTTACGAAATCCCTTTGTTCCAAAGCAACAT 146
 Db 172 TTCGGTTTGGGTATATATGCTGGGATTTTATCTACAAGCTACCAATTTCCAAACCTCAT 231
 Qy 147 TTTTCAGATTCATCGTTTCCAAACATCAACAACTCTCTCTAAAACCTCTCATACACTTC 206
 Db 232 TTCTGTTTCAACCATTTATCTCTTAAACATTTCACTCTCCCTCACTCTCAAAACACTAC 291
 Qy 207 TTCCCTCTCTCATGCAATTTTAACTTACCTTCAATCCCGGAGAAAATGCGGAGTTTCGG 266
 Db 292 ACACCTTTAGCTGGCAACATTCGTTCTCCACTAAATTCGAGTGGTTATCTGAGTTGCGT 351
 Qy 267 TATCTATCCGGGAGCTCGGTTTCTTTTCCACATGCGAGAATCTTAGCGAGCTTTTCGATGAT 326
 Db 352 TATGTCACTGGAGATCTATATCTGTTACTTTTATTGAGACTGATATGAATTTCAATCAT 411
 Qy 327 CTCGTCGGAATTCGTCAGAAATCTCCGTTAGCTCTACAACCTTTGTCCTAAATTCGCG 386
 Db 412 CTCATTTGGTAACCATCTCTGNAATGCTAAGGATTTTATCCCTTCAATTCCTCAATTTGGCA 471
 Qy 387 CCATTTGTCGAAGAATCCGATAGAAAACCTCTTCCAAAGTTTTCGCGCTGAGGTTGACTCTT 446
 Db 472 CAACCTTAAGATGACACCGGGGTCAAACTAGTCCCGGTCTTAGCCATTTCAAGTGACACTT 531
 Qy 447 TTCCAGGCGGAGCGCT---CGGTATTGGAATAGCAACGCATCACACCGTTAGCGAGCGC 503
 Db 532 TTTCGCAATCTTGGCATATCCCGTTGGTGTATTAGTAAACCATCATCTCGCTTGTGATGA 591
 Qy 504 CGCTCGTTCTTCGCGTTTATACGCGTTTGGTCTTCAATGAGCAACACATTTGAAATGAA 563
 Db 592 AATACCATCTGTAATTTCAATAGAACATGGGTTTACTCAACAAATTCGCGGTGATGAA 651
 Qy 564 GA 565
 Db 652 CA 653

RESULT 11
 BE187619 788 bp mRNA EST 22-JUN-2000
 LOCUS EST336180 KV0 Medicago truncatula cDNA clone pkV0-1613, mRNA
 DEFINITION sequence.
 ACCESSION BE187619
 VERSION BE187619.1 GI:8666803
 KEYWORDS EST.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
 1 (bases 1 to 788)
 VandenBosch,K., Endre,G., Hur,J., Moore,J., Beremand,P., Ellis,L.,
 Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and
 Fraser,C.M.
 ESTs from uninoculated seedling roots of Medicago truncatula
 Unpublished (1999)
 Contact: VandenBosch K
 Department of Biology
 Texas A&M University
 College Station, TX 77843-3258, USA
 Tel: 409 845 7707
 Fax: 409 845 2891
 Email: kate@mail.bio.tamu.edu
 Texas A&M University name:T260795e
 TIGR sequence name:MTGAU50TK
 More information is available at:
 http://chrysis.tamu.edu/medicago

```
FEATURES source
Seq primer: SKmod (CTA GAA CTA gta gAT CC) .
Location/Qualifiers
1. 788
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pkV0-1613"
/clone_lib="KV0"
/tissue_type="Seedling roots"
/dev_stage="Immediately prior to inoculation with
Sinorhizobium meliloti (0 hour)"
/note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0LR cells."
BASE COUNT 237 a 188 c 100 g 262 t
ORIGIN
Query Match 7.4%; Score 109.8; DB 164; Length 788;
Best Local Similarity 51.1%; Pred. No. 1.le-17;
Matches 311; Conservative 0; Mismatches 292; Indels 6; Gaps 2;
QY 54 TCGGTGCGCTCACATCTTCGACATGACGTGGTGCATTTTCATCCCATCGTTCAGCTC 113
DB 5 TCCCTCCCCCAACATCTTTGACATTCCTTGGCTCTTTTCTCTCCAAAGCAACCATCA 64
QY 114 CTCCTTACGAATCCCTTGTTCGAAGCAACATTTTTCAGAAATCCATCGTTCCTCAAACTC 173
DB 65 TTTTCTATGAATTCCTCAATTCATTTCCATTCACCAACACCATTTGTCCTCAAACTT 124
QY 174 AAACAATCTCTCTAAACATCTCATACACTCTCTCCCTCTCTCATGCAATTTAATCTAC 233
DB 125 AAACAATCTCTCTCTCACACTTCAACATTTACTTCCCATTTTC---CGGTACATTCGTT 181
QY 234 CTTTCATCCCGGAGAAATGCGGAGTTTCGGTATCTATCTCGGGGACTCGGTTCTTTTC 293
DB 182 CCTCGCTTGATCTAACCGAACCTCACTCGAATTTACTCTTAACACTCGGTTTCATTC 241
QY 294 ACCATCGCAGAACTPAGCGAGACTTCGATGATCTGTCGGAATTCGCCAGAAATCTCCC 353
DB 242 ACGGTCGCAGAGTCAATAGGAGCTTTGAACACTTGTCTCGGATTTACTCAAGAGATGTT 301
QY 354 GTTAGGCTCTACAACCTTGTCCCTAAATTTGCGGCCCATTTGTCGAAGATTCGATAGAAA 413
DB 302 AATGAATTCACCCCTTTGGTACCAAGTTGCAACAAA---TCTTTTCATTTTGAAGTAAA 358
QY 414 CTCCTCCAAAGTTTTCGCGCTGCAGTGACTCTTTTCCAGCGCGGCTCGGTATTGGA 473
DB 359 GAGTTCCTTTGTTAGCTATTCAATATACATCATTTCCCAATATTGCTTTTCCATTTGGC 418
QY 474 ATAGCAAGCATACACACGTTAGCGACGCCCGCTGTTCTCGGTTTATACGCGCTTGG 533
DB 419 CTAGCTTTTCATCATGTTGTTGCTGATGGAAGAACCTTCCACAAATTTTCATCAAAACATGG 478
QY 534 TCTTCATGAGCAACACATTTGAATGAAGATGAAGATGAAGATTTAAATCTTTTGCCA 593
DB 479 TCTTCATATTGTTCTCTCTCTACCTTTTGAAGACTCATCATCGCTTATAAAATCCCTTCCC 538
QY 594 GTTTTCGATAGATCGTCATAAAATATCCGACGAAATTTGATCTCATTTATTGGAGAAC 653
DB 539 TTGATGATCGAAGTGTGATTATCGACACAATGACCTTCACGAAGTTTCTTGAAGAT 598
QY 654 GCGCTAAA 662
DB 599 TGGAGAAA 607
RESULT 12
```

```
BE922746 574 bp mRNA EST 02-OCT-2000
LOCUS EST426515 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION C57B21L12 5' sequence, mRNA sequence.
ACCESSION BE922746
VERSION BE922746.1 GI:10448822
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 574)
AUTHORS Van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
UNPUBLISHED
JOURNAL Generation of ESTs from potato leaves and petioles
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES source
Location/Qualifiers
1. 574
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="C57B21L12"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 162 a 123 c 106 g 183 t
ORIGIN
Query Match 7.3%; Score 108; DB 142; Length 574;
Best Local Similarity 52.2%; Pred. No. 3e-17;
Matches 290; Conservative 0; Mismatches 260; Indels 6; Gaps 2;
QY 92 TTTTCATCCCATGCTTCAGCTCCTCTCTAGCAATTCCTTGTCCCAAGCAACATTTTTC 151
DB 3 TTTTGGCTATATGCGTCGGATTTTATTCTACAAGTACCAATTTCCAAACTCGATTTCTGT 62
QY 152 AGAATCCATCGTTCCAAACATCAAAACATCTCTCTAAAACCTCTCATACACTCTTCTCCC 211
DB 63 TCAACCATATTATTCTACTCTTTAAACATTTCACTCTCCCTCACTCTCAAAACACTACACAC 122
QY 212 TCTCTCATGCAATTTAATCTACCTTTCATCCCGGAGAAAATGCGGAGTTTCGGTATCT 271
DB 123 CTTAGCTGGCAACATTTGCTTGCCTCACTAAATTCGAGTGGTTATCTCGAGTTGCGTTATGT 182
QY 272 ATCCGGGAGCTCGGTTCTTTCACCATCGAGATCTAGCAGACACTTCGATGATCTCGT 331
DB 183 GACTGGAGATCTATATCTGTACTTTTATTGAGACTGATATGAATTTCAATCATCTCAT 242
QY 332 CGGAATCGTCCGAATCTCCGTTAGGCTCTACAACCTTTCTCCCTAAATTTGCGGCCCAT 391
DB 243 TGGTAACCATCTCGAAT---GCTAAGGATTTTATCCCAATTTCTCAATTTGGCACARACC 299
QY 392 TGTCGAAGAATCCGATAGAAAACCTTCCAAAGTTTTCGCCGTGCGAGTGTACTTTTCCC 451
DB 300 TAAGGATGACCGGGGGTCAAACTAGTCCCGGTCTTAGCCATTTCAAGTGACACTTTTTCC 359
QY 452 AGGCCGAGGCGTCGGTATTGGAATAGCAACCATCACACCGCTTAGCGACGCCCGCTGCT 511
DB 360 GAATCTTTGGCATATCCGTTGGTTTTTAGTAACCATCATGTCGCTTGTGATGAAATACCAT 419
```



```

QY 512 TCTCGGTTTATAACGGCTTGGTCTTCAATGAGCAAAACACATTGAAATGAAGA 571
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 420 CGTGAATTCATAAGAACATGGGTTTACTCAACAATTCGGCGGTGATGAACAGTCTT 479

QY 572 TGAAGAATTTAAATCTTTCCAGTTTTCGATAGATCCGTCATAAAATATCCGACGAATTT 631
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 AGAGAA---TGAGTTTCATTCATTTTATGATAGGTCGGTGATAAAAGACCCTTATGAACA 536

QY 632 TGACTCCATTTATTGG 647
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 537 AGGACGATTATATGG 552

RESULT 13
LOCUS BF096979 482 bp mRNA EST 19-OCT-2000
DEFINITION EST400516 tomato nutrient deficient roots Lycopersicon esculentum
cDNA clone cLEW18020 5' sequence, mRNA sequence.
ACCESSION BF096979
VERSION BF096979.1 GI:10902689
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 482)
AUTHORS van der Hoeven,R.S., Garvin,D.F., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T.S., Ronning,C.M., Craven,M.B., Bowman,C.L.,
Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Giovannoni,J.J.
and Tanksley,S.D.
TITLE Generation of ESTs from tomato nutrient-deficient roots
JOURNAL Unpublished (1999)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU.
FEATURES
source
1..482
/organism="Lycopersicon esculentum"
/cultivar="TA492"
/db_xref="taxon:4081"
/clone_lib="tomato nutrient deficient roots"
/clone="cLEW18020"
/tissue_type="roots"
/dev_stage="5-6 weeks old"
/lab_host="SOLR"
/note="Vector: pBluescriptSMCudapt; Site_1: 5' EcoRI;
Site_2: 3' XhoI; Roots were harvested from plants grown
under the following deficiencies/stresses: 10 mM AL, Zn,
P, K. Fe.N. mRNA was isolated from individual treatments.
Proportional aliquots of mRNA of each treatment were mixed
and used for library construction."
BASE COUNT 113 a 119 c 93 g 157 t
ORIGIN

Query Match 7.2%; Score 106.6; DB 144; Length 482;
Best Local Similarity 51.6%; Pred. No. 6.7e-17;
Matches 244; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 15 TGTAGATTTGGCGCGCGGACTCGGTGGCGGAGCAATCGGTGCGCTCACATCTTTC 74
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 TGTCAAGTTGGCGCACTCCCGCGCGGCAACGAGGTGATACCTCTTACTATTATT 63

QY 75 GACATGACGTGGTGCATTTTCATCCCATGCTTCACGCTCCTCTTTCAGCAATTCCTTGT 134
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 GACCATGTTTGGTTAGGTTTCGCCCGTATGAGCGGATATATTATTACAAGCTCTCCATT 123

```

```

QY 135 TCACAAGCAACATTTTTCAGAAATCCATCGTTTCAAAAACATCAAAACATCTCTCTAAACT 194
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 TTCAAAACCGAATTCGTTTCAAAACATTAATCTCTCTCTTAAAAAATTCATCTCTCCCTCACT 183

QY 195 CTCATACACTTCTTCCCTCTCTCATGCAATTTAATCTACCCCTTCATCCCGGAGAGAAATG 254
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 CTCAAACACATATACGCCCTTAGCCGGAAACGTTGCTTGTGCTTGTCCACTAGATACAAACGGATAT 243

QY 255 CCGGAGTTTCGGTATCTATCCGGGGACTCGGTTTCTTCCACCATCCAGAACTACGGAC 314
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 CCTGAGTTACGTTATGTGACAGGAGATTCTGTGCTGTGTTACTTTTTCGAGACTGATATG 303

QY 315 GACTTCGATGATCTCGTCGTAATCTCCAGAATCTCCCGTTAGGCTCTACAACTTTTGTG 374
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 AATTTCAATTAATCTCATCTGTCGACCATCCGCGTAAGGCTTAAGGATTTTATCACTTTGTT 363

QY 375 CCTAAATCCCGCCCATCTTCGAAGAATCCGATAGAAAACTTCCAAAGTTTTCGCGGTG 434
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 364 CCTAAGTTAGGGGAACCTTAAGGATGCACCGGGGTCCAACCTAGCCCGCTCTTAGCCATT 423

QY 435 CAGGTGACTCTTTTCCAGGCGGCGTGGTATTTGGGAATAGCAACGCATCA 487
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 CAGGTGACACTTTTTCGGAATCTTGGTGATCCATTGGTTTCTACTAACCATCA 476

RESULT 14
LOCUS BE434257 591 bp mRNA EST 24-JUL-2000
DEFINITION EST405335 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG15H6, mRNA sequence.
ACCESSION BE434257
VERSION BE434257.1 GI:9432100
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 591)
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
TITLE Generation of ESTs from tomato fruit tissue, breaker stage
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
FEATURES
source
1..591
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG15H6"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSMCudapt; Site_1: EcoRI;
Site_2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."
BASE COUNT 159 a 129 c 121 g 182 t
ORIGIN

```

```
Query Match      7.2%; Score 106.6; DB 167; Length 591;
Best Local Similarity 51.6%; Pred. No. 7e-17;
Matches 269; Conservative 0; Mismatches 249; Indels 3; Gaps 1;

Qy 102 ATGCTTACGCTCTCTTCTAGCAATTCCTTGGTTCCAAAGCAACATTTTTCAGATCCATC 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 ATGAGCGGATATTATTATACAAAGCTCTCCCATTTTCAAAACCGATTTGGTTCAAAACATT 66

Qy 162 GTTCCAAACCTCAACATCTCTCTTAACACTCTCATACACTTCTTCCCTCTCTCATGC 221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 ATTCCTCTCTTAANAATTCACCTCTCCCTCACTCTCAAAACACTATACGCCCTTAGCCGGA 126

Qy 222 AATTTAATCTACCTTCTATCCCGGAGAAAATGCGGAGTTTCGGTATCTATCCGGGGAC 281
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 AAGCTTGTCTTGCACATAGATACAAACGGATATCTCTGAGTTACGTTATGTACAGGAGAT 186

Qy 282 TCGGTTTCTTTCACCATCGCAGATCTAGCAGCACTTCGATGATCTCGTCGGAATCGT 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 TCTGTCTCTGTACTTTTTCGAGACTGATATGAATTTCAATTTATCTCATTTGGTGACCAT 246

Qy 342 CCAGAACTCCCGTTAGGCTCTACAACTTTCCTCTAAATTTGCCGCCCATTTGTGGAAGAA 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 CCGGTAGGCTTAGGATTTTATACACTTTTGTCTTAAGTTAGGGGAACCTAAGGATGCA 306

Qy 402 TCCGATAGAAAACCTTCCAAAGTTTTCGCCGTGCGAGTGACTCTTTTCCCGAGCGGAGGC 461
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 CCGGGGTCACACTAGCCCGCTCTTACCCATTCAGGTGACACTTTTTCGGAATCTTGGT 366

Qy 462 GTCGGTATTGGAATAGCAACGATCAGACCGTTAGCGAGCGCCCGCTGTTCTCGCGTTT 521
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 GTATCCATTTGTTTTCACTAACCATCATCTTGTGTGATGAGCTACTATAGCAGGGTTC 426

Qy 522 ATAAGCGCTTGGCTTCTCAATGAGCAACACATTTGAAATGAAGATGAGATGAAGATTT 581
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 ATTAAGCGGTGGGCTCTACTCCACAAATTCGGTGGACATGAA---CAATTTCTATCGAAT 483

Qy 582 AAATCTTTGCCAGTTTTCGATAGATCCGTCATAAAATATCC 622
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 484 GAGCTAATTCATTTATGATAGTGGTCCGTAGTAAAGACCC 524

RESULT 15
AW616206      487 bp mRNA EST 24-MAR-2000
LOCUS      EST307245 L. hirsutum trichome, Cornell University Lycopersicon
DEFINITION      hirsutum cDNA clone cLHTID15 5', mRNA sequence.
ACCESSION      AW616206.1 GI:7322240
VERSION      EST.
KEYWORDS      EST.
SOURCE      Lycopersicon hirsutum.
ORGANISM      Lycopersicon hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE      1 (bases 1 to 487)
AUTHORS      van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang
,F., Hansen,T., Craven,M.B., Bowman,C.L., Konning,C.M., Nierman,W.,
Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from wild tomato (Lycopersicon hirsutum)
trichomes
JOURNAL      Unpublished (2000)
COMMENT      Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
FEATURES      Location/Qualifiers
source      1..487
```

```
/organism="Lycopersicon hirsutum"
/db_xref="taxon:62890"
/clone_lib="L. hirsutum trichome, Cornell University"
/dev_stage="mixed stages"
/notes="Leaves of various stages were shaken in liquid
nitrogen, shearing off trichomes. This procedure yielded a
mixture of cells which is highly enriched for trichome
likely with minor contaminations of other types of leaf
cells."
BASE COUNT      116 a      117 c      91 g      163 t
ORIGIN
```

```
Query Match      7.1%; Score 104.8; DB 118; Length 487;
Best Local Similarity 51.3%; Pred. No. 2e-16;
Matches 244; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

Qy 26 GCCGCCGCCGACTCGGTGGCGGAGCAATCGGTCCGCTCACATTTCTTTCGACATGACGTG 85
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 GCCACCTCCCAAGCGCGCAACGAGGTGGTACTTCCCTTACTTATTTGACCATGTTTG 60

Qy 86 GCTGATTTTCATCCCATGCTTCCAAACACTCAAACTCTCTCTAAACTCTCATACATT 145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 GTTAGGGTTTCCACCTTATTAGGCGGATATTATTCTACAAGCTCTCCATTTCCAAATCCGA 120

Qy 146 TTTTTCAGAAATCCATGCTTCCAAACACTCAAACTCTCTCTAAACTCTCATACATT 205
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 TTTCTGTTCAAAACATTAATTCCTCTCTTAAAAATTCACATTTTCCCTCACTCTCAAAACATA 180

Qy 206 CTTCCCTCTCTCATGCAATTTAAATCTACCCCTTCATCCCGGAGAAAATGCCGAGTTTCG 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 TATGCCCTTAGCCGGAACGTTGCTTGTCCACTAGATACAAACGGATATCTGAGTTAGC 240

Qy 266 GTATCTATCCGGGAGCTCGGTTTCTTTCACCATCGCAATCTAGCGAGACTTTCGATGA 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 TTATGTGACAGGAGATCTGTGTCTGTTAAATTTTCTGAGACTGATATGGATTTCAATTA 300

Qy 326 TCTCGTCGGAATCGTCCAGAAATCTCCCGTTAGGCTCTAGCACTTTGCCCTTAATGCC 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TCTCATTGGCGACCATCCGCTAATGCTAAGGATTTTATCATTGTTTCTCAAGTTAGG 360

Qy 386 GCCCATTTGTCGAAGAATCCGATAGAAAACCTCTTCCAACTTTTCCGCCGTCAGGTGACTCT 445
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 GGAACCTAAGGATGCACCCGGGTCCAAATTAGCCCGCTCTTAGCCATTCAAGTGACACT 420

Qy 446 TTTCCCGAGCGGAGCGGTGGTATTGGAAATAGCAACGCATCACACCGTTAGCGAGC 501
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 TTTTCCGAATCTTGGTGTATCCATTGGTTTTCACCTAACCATCATGTTGTTGGTGATG 476
```

Search completed: November 5, 2001, 18:00:56
Job time: 14896 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 16:50:26 ; Search time 8904.87 Seconds
(without alignments)
2619.394 Million cell updates/sec

Title: US-08-894-356C-5
Perfect score: 1508
Sequence: 1 TGAACATTCGACATGCC.....AAAAAAAAAAAAAAAAAAAA 1508

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues 2688314
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_om.*
8: gb_ov.*
9: gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_pl1.*
13: gb_pl2.*
14: gb_pl3.*
15: gb_pl4.*
16: em_ba1.*
17: em_ba2.*
18: em_fun.*
19: em_htgo_hum.*
20: em_htgo_inv.*
21: em_htgo_rod.*
22: em_htg_hum1.*
23: em_htg_hum2.*
24: em_htg_hum3.*
25: em_htg_hum4.*
26: em_htg_hum5.*
27: em_htg_hum6.*
28: em_htg_hum7.*
29: em_htg_hum8.*
30: em_htg_inv1.*
31: em_htg_inv2.*
32: em_htg_other.*
33: em_htg_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pl.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vl.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vil.*
59: gb_vl2.*
60: gb_htg1.*
61: gb_htg2.*
62: gb_htg3.*
63: gb_htg4.*
64: gb_htg5.*
65: gb_htg6.*
66: gb_htg7.*
67: gb_htg8.*
68: gb_htg9.*
69: gb_htg10.*
70: gb_htg11.*
71: gb_htg12.*
72: gb_htg13.*
73: gb_htg14.*
74: gb_htg15.*
75: gb_htg16.*
76: gb_htg17.*
77: gb_htg18.*
78: gb_htg19.*
79: gb_htg20.*
80: gb_htg21.*
81: gb_htg22.*
82: gb_htg23.*
83: gb_htg24.*
84: gb_htg25.*
85: gb_pr1.*
86: gb_pr2.*
87: gb_pr3.*
88: gb_pr4.*
89: gb_pr5.*
90: gb_pr6.*
91: gb_pr7.*
92: gb_pr8.*
93: gb_pr9.*
94: gb_r01.*
95: gb_r02.*
96: gb_in4.*
97: gb_pr10.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DE	ID	Description
1	1501.6	99.6	1508	10	E12757	E12757 Senecio cru
2	202.4	13.4	1679	12	AB010708	AB010708 Gentiana
3	202.4	13.4	1703	10	E12753	E12753 Gentianatri
4	189.2	12.5	1476	12	AB029340	AB029340 Perilla f
5	185.8	12.3	1479	10	E12756	E12756 Perilla oci
6	138.2	9.2	1622	10	E12754	E12754 Gentianatri
7	138.2	9.2	1622	12	AB026494	AB026494 Gentiana
8	114.2	7.6	101176	12	AC002560	AC002560 Genomic s

QY 841 GCATAGCAATCACTCGTAAATAGGAGAAAGGCGAAGACGAGTTAGAACACT 900
|||||
Db 841 GCATAGCAATCACTCGTAAATAGGAGAAAGGCGAAGACGAGTTAGAACACT 900
QY 901 TCATATCAACCAATGATTCGATCGCTGCTGATGACCAATTCACGAGCTACTTTG 960
Db 901 TCATATCAACCAATGATTCGATCGCTGCTGATGACCAATTCACGAGCTACTTTG 960
QY 961 GTAACGTGGTGCACCATGTCGCGACCTTAAATAATGCTTTTGACTACGGAATG 1020
|||||
Db 961 GTAACGTGGTGCACCATGTCGCGACCTTAAATAATGCTTTTGACTACGGAATG 1020
QY 1021 GGTATGCACCTTGGTCTAAAGTAATGGAGAGCTATATGCAAAATGATATAAAGG 1080
|||||
Db 1021 GGTATGCACCTTGGTCTAAAGTAATGGAGAGCTATATGCAAAATGATATAAAGG 1080
QY 1081 ACGGAATCTTGAAGATGCGCGAGATGCGATGAACCTTTTCATGATCCCGGTAGGAAGA 1140
Db 1081 ACGGAATCTTGAAGATGCGCGAGATGCGATGAACCTTTTCATGATCCCGGTAGGAAGA 1140
QY 1141 TTGGTGTCTGCTACACCTAAGCTCAACTTGTACGACTTTGATTTGGTGGGGAAGC 1200
|||||
Db 1141 TTGGTGTCTGCTACACCTAAGCTCAACTTGTACGACTTTGATTTGGTGGGGAAGC 1200
QY 1201 GCATAAAGTATGAGACTGTTTCAATAGACTATAATACGTCGATTTCTATAAATGCAAGCA 1260
|||||
Db 1201 GCATAAAGTATGAGACTGTTTCAATAGACTATAATACGTCGATTTCTATAAATGCAAGCA 1260
QY 1261 AAACATCAGCACCAAGATCTTGAATGGATGAGTCTACCGAGTATGCAAAATGGAGCGT 1320
Db 1261 AAACATCAGCACCAAGATCTTGAATGGATGAGTCTACCGAGTATGCAAAATGGAGCGT 1320
QY 1321 TTTCATGATCTTTGATGAAGATGATGAGAGTCAAGTTTCATGATGATCATCGTCCCT 1380
Db 1321 TTTCATGATCTTTGATGAAGATGATGAGAGTCAAGTTTCATGATGATCATCGTCCCT 1380
QY 1381 TTTCATGATCTTTGATGAAGATGATGAGAGTCAAGTTTCATGATGATCATCGTCCCT 1440
Db 1381 TTTCATGATCTTTGATGAAGATGATGAGAGTCAAGTTTCATGATGATCATCGTCCCT 1440
QY 1441 TACCTTTCGACTATGTTTGAAGATGCAACGATATAAATGAAAAAAGAAAAA 1500
Db 1441 TACCTTTCGACTATGTTTGAAGATGCAACGATATAAATGAAAAAAGAAAAA 1500
QY 1501 AAAAAA 1508
Db 1501 AAAAAA 1508

RESULT 2
LOCUS AB010708
DEFINITION Gentiana triflora mRNA for Anthocyanin 5-aromatic acyltransferase,
complete cds.
ACCESSION AB010708
VERSION AB010708.1 GI:4185598
KEYWORDS Anthocyanin 5-aromatic acyltransferase.
SOURCE Gentiana triflora petal cDNA to mRNA, clone: pGAT4.
ORGANISM Gentiana triflora
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Gentianales; Gentianaceae; Gentiana.
1 (bases 1 to 1679)
REFERENCE
AUTHORS Fujiwara, H., Tanaka, Y. and Kusumi, T.
TITLE Direct Submission
JOURNAL Submitted (22-JAN-1998) to the DDBJ/EMBL/GenBank databases.
Hirotsuki Fujiwara, Suntory LTD., Inst. of Fundamental Research;
1-1-1, Wakayamadai, Yamamoto-cho, Osaka 618-0024, Japan
(E-mail: Hirotsuki.Fujiwara@suntory.co.jp, Tel: +81-75-962-8807,
Fax: +81-75-962-8262)
2 (sites)
REFERENCE
AUTHORS Fujiwara, H., Tanaka, Y., Yonekura-Sakakibara, K.,

TITLE
JOURNAL
MEDLINE
FEATURES
source

Fukuchi-Mizutani, M., Nakao, M., Fukui, Y., Yamaguchi, M., Ashikari, T.
and Kusumi, T.
cDNA cloning, gene expression and subcellular localization of
anthocyanin 5-aromatic acyltransferase from Gentiana triflora
Plant J. 16 (4), 421-431 (1998)

Location/Qualifiers
1. .1679

/organism="Gentiana triflora"
/db_xref="taxon:55190"
/clone="pGAT4"
/tissue_type="petal"
6. .1415
/codon_start=1
/product="Anthocyanin 5-aromatic acyltransferase"
/protein_id="BAA74428.1"
/db_xref="GI:4185598"
/translation="MEQIQMKVLEKQVTPSPDTPVELSLPVTFFDPMHLNKM
SLFYDFPYPRTHFDVTPNLKASLSLTHKHYPLSGNLLMPKSGMPKQFQSRDE
GDSITLIVAESDQDFYLLKQHLVDSNDLHGLFYVMPRTVMQYKVIPLAVQVTV
FNRGIAVALTAHHSIADAKSFVMPINAWAYINKFGKDALLSANLLSFDRSIIKDL
YGLETFWNEHQDVLEMFSGRSPKPPREKVRATYVLSLAELQKLNKLVNLRGSEPT
IRVITFTMTGVTWCMKSDVVSESSNDNELEYFSETADCRGLTTPCPNPF
GNCLASCVAKXATHEKLVGDLVAVAAIGEAIEKRLHNEKGVLDADKTWSESGIP
SKRFLGTGSPKFDYGVDFGKPAKFDITSDVYAEIYVQSRDFEKGVEIGVSLP
KIHMDAPAKIFEFGFCSLS"

polya_site

/note="24 a nucleotides"

BASE COUNT 488 a 350 c 359 g 482 t

ORIGIN

Query Match 13.4%; Score 202.4; DB 12; Length 1679;
Best Local Similarity 52.0%; Pred. No. 1.5e-34;
Matches 720; Conservative 0; Mismatches 616; Indels 48; Gaps 10;

QY 1 TGAACATCTCTGCAACATGCCGAATATCGGCCCTCGGGCACCA---TCGGCCATCGCT 57
Db 25 TGAAGTCTCTGCAAAATGCCAAGTTACACCACCATCTGCACACAGATGCGAGTTAT 84
QY 58 CGTTATCTCTTACTTTCTGACATTTACCTGCTACTTTCCCTCGGGTCCACCATCTTTT 117
Db 85 CGTACCGGTAAACATCTTTCGATATCCCTGTTGCATTTGAATAAGATGCGAGTCTTC 144
QY 118 TCTTCTATCATCTTCCACATTTCAATCCATTTCTATGACGACACTATTTGTCACGCTAA 177
Db 145 TGTTTTACGACTTTTCGTACCCCAAGAACACATTTCTTTGGACACTGTTATCCCTTAATCTTA 204
QY 178 AACAACTTTTATCGGTCACTCTTCAACATTTTTCCTGCTAGTAAATTTGATTTGAT 237
Db 205 AGGCTCTTTGCTCTCACTCTAAACACTAGCTTCCGCTTAGCGGAATTTGTTGATGC 264
QY 238 TTCTTAACACTGATGGTTCGGGTTTTTAATAAAACACAGAAATAAACACGTTGAAGGTG 297
Db 265 CGATCAA-----ATCGGCGAAATGCCAAGTTTCAGTACTCCCTGATGAGGGCG 315
QY 298 ATTCTGTTGTTGTTACTTTTGCAGATGTTGCTGACTTTAATAATTTGACAGGAATC 357
Db 316 ACTCGATAACTTTGATGCTGTCGGAGTCTGACCAAGGATTTGACTACCTTAAAGGATC 375
QY 358 ATCCCTCGAAATGTGAAAACTTTTATCCACTTGTACTCTTTCATTTGGGAAATGCAATCAAT 417
Db 376 AACTGGTAGATTCATGATTTGCCATGGCTTTTATGTTATGTCACCGGTTATTAAGA 435
QY 418 ----TATGTTGATGCGTACGGTCCCACTTTTTCACCTTCAAGTGACGTTTTCGCGGCT 474
Db 436 CCATGCAAGACTATAAAGTGATCCCGCTGTCAGCCGTCGAAGTAAACCGTTTTCCTAAC 495
QY 475 CGGGTATATCACTAGTAATGACCAATCATCATAGCTTTGTTGACGCTACGACCGGTTCA 534
Db 496 GTGGCATGCCGCTGCTGACGGGACATCATCAATTCAGATGCTAAAGGTTTGTAA 555
QY 535 ACTTTTGAAGGGGTGGACTTCGATTATTCAATCTGCTGATGATCGTCTCTTTTAAACGA 594

Db	556	TGTTTCATCAATGCTTGGGCCCTATATTACAAATTTGGGAAGACGGGACTGTTGTCCG	615
Qy	595	AAGGATCTCCACCGGTTTTCATAGATTGATTAACA-----TCCACATTTAGATGA	646
Db	616	CGAATCTTCTCCATCTTTCGATAGATCGATTAACAAGATCTGTATGGCCCTAGAGGAA	675
Qy	647	AAATAAGTTGAGACATACAAG---GCTCGAAAGTTTTTATAAACCTTCGAGCCTGTTTG	702
Db	676	CATTTTGGGAACGAATGCAAGATGTTCTTGAATGTTCTCTAGATTTGGAAGCAACCCC	735
Qy	703	GTCCCACTGATAAAGTTCGGTCAAGTTTGTGTGACCCGAACTAAATCAATCTACTAA	762
Db	736	CTCGATTCAACAAGGTACGAGTACATATGTCCTCCTCCCTTGCCTGAATCCAGAAGCTAA	795
Qy	763	AGAAAAGGCTTAA---CCCAAGTGCCAACTTGGAGTACATGT--CATCTTTTACGG	816
Db	796	AGACAAGTACTGAATCTCAGAGATCCGAACCGACAAATACGTGTACGACGTTTCACAA	855
Qy	817	TAACCTTGTGTTATATATGAGTTGTCATAGCGAAATCA-----CTCGTAAAAATAG	867
Db	856	TGCGTGTGGATACGTATGAGATGTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	915
Qy	868	GAGAAAGAAAGGGGAGACGAGTGAACAGTTTCAATCAACCATGATGTTGTCGATCTC	927
Db	916	AATCATCGAACGACGAAATGAGCTCGAGTACTTTCAGTTTTCACAGCGGATTCGCGGAC	975
Qy	928	GTCCTTGATCCACCAATCCCAACGCTACTTGTGTAAGTGGTGCACCATGTCTCCGA	987
Db	976	TTCTGACCCCCCGTCCGCTTAACTTTCGCACTTTCGCACTGTCCTTTCGCGTCATGCGA	1035
Qy	988	CCTTAAAAATGTCGTTTGTGACTACGGAATGGGTATGCACTTGGTGGCTAAAGTAATTG	1047
Db	1036	AAGCAACACATAAAGAGTTAGTTGGGATAAAGGCTTCTTTGTTGCAGTTTCAGCTATTG	1095
Qy	1048	GAGAGTCTATATGCAAAATGATATATAAGAGCGGAATCTTGAAGATGCCGCGAGAT	1107
Db	1096	GAGAAGCCATTGAAAGAGGTTGACACAGAAAAGGCGTTCTTTCGAGATGCAAAAACCT	1155
Qy	1108	GG---CATGACCTTTCATGATCCGGCTAGGAAGATT---GGTGTGCTGCTACACCTA	1161
Db	1156	GGTATCGGAATCTAATGGAATCCCTTCAAAAAGATTTCTCGGGATACCGGATCGCCTA	1215
Qy	1162	AGCTCAACTGTACGACTTTGATTTTGGTGGGGGAGCGGATCAAGTATGAGACTGTTT	1221
Db	1216	AGTTCGATTCTGATGGTGTAGATTTTGGATGGGAAAGCCTGCAAAATTTGACATTACCT	1275
Qy	1222	CAATAGACTATAATACGTCGATTTCTATAAATGCAAGCAAAACATCAGCACAAATCTTG	1281
Db	1276	CTGTGATTATGCGAATTTGATTTATGCTGATTCAGTCCAGGGATTTTGA AAAAGGTGG	1335
Qy	1282	AAATTGGATTGAGTCTACCGAGTATGCAAAATGAGGGGTTTTTCTAGGATCTTTGATGAAG	1341
Db	1336	AGATTGGAGTATCATTCGCTAAGATTTATATGGATGATTTTGCAAAAATCTTTGAAGAAG	1395
Qy	1342	GATT 1345	
Db	1396	GCTT 1399	

RESULT 3

E12753					
LOCUS	E12753	1703 bp	DNA	PAT	24-JUN-1998
DEFINITION	Gentianatriflora mRNA acyltransferase, complete cds.				
ACCESSION	E12753				
VERSION	E12753.1	GI:3251585			
KEYWORDS	JP 1997070290-A/1.				
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1703)				
AUTHORS	Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y., Yonekura, K., Mizutani, M. and Kusumi, T.				

TITLE		GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY	
JOURNAL		Patent: JP 1997070290-A 1 18-MAR-1997;	
COMMENT		SUNTORY LTD	
		OS Gentianatriflora	
		PN JP 1997070290-A/1	
		PD 18-MAR-1997	
		PF 30-JAN-1996 JP 1996046534	
		PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI	
		ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO	
		MASAHITO.	
		PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAAKI PC	
		C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC	
		C12R1:865),	
		PC (C12N9/10,C12R1:19);	
		CC strandedness: Double;	
		CC topology: Linear;	
		FH Key Location/Qualifiers	
		FT source 1..1703	
		FT /organism='Gentianatriflora'	
		FT /variety='japonica'	
		FT /tissue_type='petal'	
		FT /clone='pGAT4'	
		FT CDS 6..1415	
		FT /product='acyltransferase precursor' FT	
		mat_peptide 9..1412	
		FT /product='acyltransferase'.	
FEATURES		Location/Qualifiers	
source		1..1703	
		/organism="unidentified"	
		/db_xref="taxon:3264"	
BASE COUNT		512 a 350 c 359 g 482 t	
ORIGIN			
Query Match		13.4%; Score 202.4; DB 10; Length 1703;	
Best Local Similarity		52.0%; Pred. No. 1.5e-34;	
Matches 720; Conservative		0; Mismatches 616; Indels 48; Gaps 10;	
Qy	1	TGAACATCTCGAACATGCCGGAATATCGGCCCTCGGGCACCA--TCGGCCATCCGT	57
Db	25	TGAAGGTTCTTGAATAATGCCAAGTTACACCACCTCGACACACAGATGTCGAGTTAT	84
Qy	58	CGTTATCTTCTACTTTCTTCGACATTTACTTGGCTACTCTTCCCTCCGGTCCACCATCTTT	117
Db	85	GGCTACCGGTAACATCTTCGATATCCCTCGTTCGACATTCGAATAAGATGCACTCCCTTC	144
Qy	118	TCCTTATGACTTTCACATTTCTAAATCCCATTTTCATCGAGACACTATTTGCCAGGCTAA	177
Db	145	TGTTTACGACTTTCCGTACCCAAAGAACACATTTCTTGGACACTGTTATCCCTAATCTTA	204
Qy	178	AACAATCTTTATCGTCACCTTTCACACATTTTTCGCCGTTTGTAGTAATTTGATTGTAT	237
Db	205	AGGCTCTTTGCTCTCCTACTCTAAACACACTACGTTCCGCTTAGCGGAAATTTGTTGATGC	264
Qy	238	TTCTAAACACTGATGTTTCGGGTTTTTAATAAAACACAGAAATAAACACCGTTGAAGGTG	297
Db	265	CGATCAA-----ATCGGGCGAATGCCGAAGTTTCAGTACTCCGCTGATGAGGCG	315
Qy	298	ATCTGTGTGGTTACTTTTGCAGAATGTTGTCTTGTACTTTTAATAATTTGACAGGAATC	357
Db	316	ACTCGATAACTTTCATCGTTCGGGAGTCTGACCAGGATTTTGGACTACCTTAAAGGTGATC	375
Qy	358	ATCCTCGAAATGTAACATTTTATCCACTTTGACCTTCACTTTCATTCGGGAATGCAATCAAT	417
Db	376	AACTGGTAGATTCCAATGATTGATGCCCTTTTATGTTATGCCACGGGTTTAAAGGA	435
Qy	418	---TATGTGATTGGTCACGCTCCACTTTTTCACTTCAAGTGACGTTTTTTCGGGCT	474
Db	436	CCATGCAAGACTATAAAGTGATGCCGCTCGTAGCGGTCGAAGTAAACGTTTTTCTTAACC	495
Qy	475	CGGCTATATCACTAGGAATGACGAATCATCATAGCCTTGGTGTAGCGTAGCACGCGGTTCA	534

Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.
1 (sites)
Sakakibara,K.Y., Tanaka,Y., Mizutani,M.F., Fujiwara,H., Fukui,Y., Ashikari,T., Yamaguchi,M. and Kusumi,T.
Molecular and biochemical characterization of a novel hydroxycinnamoyl-CoA: anthocyanin 3-O-glucoside-6
Unpublished (1999)
2 (bases 1 to 1476)
Sakakibara,K.Y. and Tanaka,Y.
Direct Submission
Submitted (24-JUN-1999) to the DDBJ/EMBL/GenBank databases. Keiko Y Sakakibara, Sumoto-y Research Center, Fundamental Research, Plant Biotech; Shimamoto-cho, Wakayamadai, 1-1-1, Mishima-gun, Osaka 618-8503, Japan (E-mail:Keiko.Sakakibara@suntory.co.jp, Tel:+81-75-962-8807, Fax:+81-75-962-8262)
Location/Qualifiers
source
/organism="Perilla frutescens"
/db_xref="taxon:48386"
/tissue_type="leaf"
<1..1343>
/codon_start=3
/product="anthocyanin acyltransferase"
/protein_id="BAAG3475.1"
/db_xref="GI:7415646"
translation="VIEPTCRVPDPSVAQSVPLTFEDMTLWLFHFMLOLLFLFEPKESKHQSESIIVPKLKQSLKTLHFELPISCNLIYSSPEKMPERYLSGDSVSSTIAESSDFDLGNRPESPVRILNFKPIVEESDRKLQFAVQTLPFGRGVGIGIAT HHTVDADPEFLAFITAMSSMKHIENEDDEEFKSLPVFORSVIKTKFDYSVIWRNAA LKFQLSDREPLTDRIITFTVFQTSKIKKLKGIQSRVPSLVHLFSFVAIAAYMKG ITKSATEDDQNDADFLIPDLRPLRPDPENYPFNCLSYALPMRRRELVLGVKGF VFLAAEVAAEIKKIRINDKRILETVKEKSPERIKALOKSYFSVAGSKLDLYGDGFQ GKARQETLSIDGKYAMTLCARDFEGGLEVCLSLKPDKMDAFAPAFVSLING"
BASE COUNT 419 a 315 c 331 g 411 t
ORIGIN
Query Match 12.5%; Score 189.2; DB 12; Length 1476; Best Local Similarity 48.5%; Pred. No. 1.2e-31; Matches 725; Conservative 0; Mismatches 738; Indels 33; Gaps 6;
Qy 5 CATTTCTCGAACATGCCGCATATCGGCCCCCTCGGGCACCATCGGCCCATCGCTCGTTATC 64 Db Qy 65 TCTTACTTTTCGACATTACTTTGGCTACTCTTCCCTCGGGTCCACCACATCTTTTCTTCTA 124 Db Qy 125 TGACTTTCACATTCATAATCCATTTTCATGACACACTATTTGTCCCAGGCTTAACAATC 184 Db Qy 185 TTTATCGGTCACTCTTCAACATTTTTTCCCGTTTGCTAGTAGTAATTTGATTGTATTCCTAA 244 Db Qy 245 CACTGATGTTGGGTTTTTAATAAAAAACCAAAAATAAAACACGTTGAAGGTGATCTCGT 304 Db Qy 305 TGTGGTTACTTTTGCAGAATTTGCTTGACITTAAATATTTGACAGGAATCATCTCTCG 364 Db Qy 365 AAAATGTGAACACTTTTATCCACTTGTACTCTTATTGGGAATGCAATCAAAATATGTGA 424 Db Qy 425 TTGCGTCACGGTCCCACTTTTTCACCTTCAAGTGACGTTTTTTCGGGCTCGGGTATATC 484 Db
RESULT 4
AB029340
LOCUS AB029340 1476 bp mRNA PLN 04-APR-2000
DEFINITION Perilla frutescens mRNA for anthocyanin acyltransferase, partial cds.
ACCESSION AB029340
VERSION AB029340.1 GI:7415645
KEYWORDS anthocyanin acyltransferase.
SOURCE Perilla frutescens leaf cDNA to mRNA.
ORGANISM Perilla frutescens

Db 407 TAGAAACCTCTCCAAAGTTTTCGCCGCTGAGAGTGACTCTTTTCCAGCGCGAGCGGTGG 466
Qy 485 ACTAGGAATGACGAATCATATAGCTTGGTGACGCTAGCAGCGGTTCACACTTTTGTAA 544
Db 467 TATTGGAATAGCAACGCATCACACCGTTAGCAGCCCGCTGTTCTCGCTTTTATAAC 526
Qy 545 AGGGTGGACTTCGATTATTAACATCTGGTGTAGATCGGTCTTTTAAAGAAAG---ATC 601
Db 527 GGCTTGGTCTTCAATGAGCAACACATTAAGAAATGAAGATGAAGAAATTTAAATC 586
Qy 602 TCCACCGTTTTTGTAGAT---TGATTAACTCCACATTTAGATGAATAAGTTGAG 658
Db 587 TTTGCCAGTTTTCGATAGATCGGTATATAAATATCGAGAAATTTGACTCCATTATTG 646
Qy 659 ACATACAGGCTCGAAAGTTTTTATAAACCTTCGAGCCTTGTGTCCCACTGATAAAGT 718
Db 647 GAGAACGCGCTAAATTTCCCTTTCGATCTCGTATCCCTCATTTACCGACGGACCGCAT 706
Qy 719 TCGGTCAACGTTTGTGTGACCGCACTTAATCAATCTACTTAAGAAAGGCTTTAAAC 778
Db 707 TCGAACCAACGTTTGTGACCGCACTTAATCAATCTACTTAAGAAAGGCTTTGATTCAGTC 766
Qy 779 CCAAGTGCACAACTTGGAGTACATCTCTTTTACGCTAACTTGTGCTTATATATGAG 838
Db 767 CAGAGTCCCAAGTTTGTGCT 826
Qy 839 TTGCATAGCGAAATCACCTCGTAAATAATAGGAGAAAGGCGGAGAGCTTTAGAACA 898
Db 827 TGGCATACGAATCATTC-----ACAGCAGATGAAGCAACAGACAGGAGTGCATT 880
Qy 899 GTTCAATATACCAATGATGTGATCTGCTTGTGATCCACCAATTTCCACAGCCTACTT 958
Db 881 TTTCTGTGATTCGGTCTGATTAAGGCCAGATTAAGTCCGCGGTTCTGAAATTAATT 940
Qy 959 TGGTAACCTGTGGTGCACCATGTGTCACCACTTAAATAATGTCGTTTGTACTACGAAA 1018
Db 941 CGGGAAGCTGCTTATCGTACGCGCTCGCGAGAAATGCGGCGGAGAGCTGGTGGGAGAA 1000
Qy 1019 TGGGTATGACCTTGGTCTAAAGTAATTTGAGAGTCTATATGCAAAATGATATATAATA 1078
Db 1001 AGGGTGTCTTGGCAGCTGAGGTAAATGCGCGGAGATTAATAAAGGATCAACGACAA 1060
Qy 1079 GGACGGAATCTTGAAGATGCGCGAGATGGCATGA---ACCTTTTCATGATCCCGGTAG 1135
Db 1061 GAGAAATATTAGAAACGTTGAGAAATGCTGCGCGAGATTCGTAAGCGTTGACAGAAATC 1120
Qy 1136 GAAGATTGGTGTGCTGTACACCTAACTCAACTTGTACACTTTGATTTTGGTGGGG 1195
Db 1121 ATATTTTTCGGTGGCAGATCGAGCAAGCTAGATCTTTACGGTGCAGATTTTGGATGGG 1180
Qy 1196 GAAGCGCATAAAGTATGAGACTGTTTCAATAGA---CTATAATACGTCGATTTCTATATA 1252
Db 1181 GAAGCGAAGCAAGCAAAATATTGTCGATTGATGGGAGAAATATGCAATGACACTTTG 1240
Qy 1253 TCGAAGCAAAACATCAGCACAAAGATCTTGAATTTGATTTGAGTCTTACCGAGTATGCAAT 1312
Db 1241 TAAAGCCAGGATTTGGAAGGAGGATTTGAGTTGCTTGTCTTTCGCTTAAGGACAAAT 1300
Qy 1313 GGAGCGTTTCTAGCATCTTTGATGAGGATAGAGAGTCAAGTTTCATTTGTAGATCAT 1372
Db 1301 GGATGCTTTTGTGCTTATTTTACCTGGGAATTAAGTTTAATAAATGATGATATAA 1360
Qy 1373 CGTCCCTTTTGTGTCATCAAGTTTCTGCTGTTTTTATGAGTTGCCACTGTTCTATTTC 1432
Db 1361 ACTAATATTGTTGTACAAATTAATTAAGTTTTCAGTAACTGAAGATATATACCTATTA 1420
Qy 1433 TTTAAGTATACCTTTCGACTATGTTTGAAGATGCAACGATATATAAATGAAAAA 1488
Db 1421 TATATTTATGATTTGGTTCAATAAAGTTAAAGCCTCTTGAAGAAAAA 1476

RESULT
E12756

LOCUS E12756 1479 bp DNA PAT 24-JUN-1998
DEFINITION Perilla ocimoides mRNA for acyltransferase,partial cds.
ACCESSION E12756
VERSION E12756.1 GI:3251588
KEYWORDS JP 1997070290-A/4.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1479)
AUTHORS Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y.,
Yonekura,K., Mizutani,M. and Kusumi,T..
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 4 18-MAR-1997;
SUNTORY LTD
COMMENT OS Perilla ocimoides
PN JP 1997070290-A/4
PD 18-MAR-1997
PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO,
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAAKI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,C12N9/10, PC
C12R1:865),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FH source 1..1479
FT /tissue_type='leaves',
FT CDS 1..1343
FT /product='acyltransferase'.
FEATURES
source Location/Qualifiers
1..1479
/db_xref='taxon:32644'
BASE COUNT 420 a 316 c 331 g 411 t 1 others
ORIGIN

Query Match 12.3%; Score 185.8; DB 10; Length 1479;
Best Local Similarity 49.4%; Pred. No. 6.7e-31;
Matches 673; Conservative 0; Mismatches 657; Indels 33; Gaps 6;

Qy 5 CATTCTCGAATCGCCGAATATCGGCCCTCGGGCCACCATCGGCCATCGCTCGTTATC 64
Db 2 CGTGATCGAAACGTTAGAGTTGGCGCGCGGACTCGGTGGCGGAGCAATCGTGCC 61
Qy 65 TCTTACTTCTTCGACATTACTTGGCTACTCTTCCCTCGGTCCACCATCTTTCTTCTTA 124
Db 62 GCTCACAATCTTCGACATGAGCTGGCTGCATTTTCATCCCATGCTTCAGCTCCTCTTA 121
Qy 125 TGACTTTCCACATCTTAATCCCATTTTCATGGACACTATTGTTCCAGGCTAAAAACAATC 184
Db 122 CGAATTCCTTGTCCAGCAACATTTTTCAGATCCCATCGTTCCAAACATCAACAATC 181
Qy 185 TTTATCGGTCTACTTTCACCAATTTTTTCCCGTTTGTCTAGTAATTTGATTTGTTCCFAA 244
Db 182 TCTCTCTAAAACTCTCATACACTTCTTCCCTCTCTCATGCAATTTTAATCTACCCCTTCATC 241
Qy 245 CACTCATGTTTCGGGTTTTTAATAAAAAACCAAGATAAAACACGTTGAAGGTGATCTGT 304
Db 242 CCCGG-----AGAAAAATGCCGAGTTTCGGTATCTATCCGGGACTCGGT 286
Qy 305 TGTGTTACTTTTGCAGAAATGTTGTTCTGACTTTTAATAATTTGACAGGAAATCATCCCTCG 364
Db 287 TTCCTTCCACATCGCAGAAATCTAGCGAGACTTCGATGATCTCTCGGAAATCTGCCAGA 346
Qy 365 AAAATGTGAAACACTTTTATCCACTTGTACCTTCATTTGGGAAATGCAATTAATATGTA 424

Db 347 ATCTCCGTTAGGCTCTACAACATTTGTCCCTAAATTTGCCGCCCATTTGTCGAAGAATCCGA 406
QY 425 TTGGCTCAGGTCCTCCACTTTTTCACCTTCAAGTGACGTTTTCCTGGGCTCGGGTATATC 484
Db 407 TAGAAACACTCTTCCAAGTTTTCCTGGTGCAGGTGACTCTTTTCCAGCGCGAGCGTCGG 466
QY 485 ACTAGGAATCAGCAATCATAGCTTGGTGACGCTAGCACCGGTTCAACTTTTGTAA 544
Db 467 TATTGGAATAGCAACGATCACACCGTTAGCGAGCGCCGCTGTTCTCGCTTTATAAC 526
QY 545 AGGGTGACTTCGATTAATCAATCTCGGTAGATCGGCTCTTTTAAACGAAAGG---ATC 601
Db 527 GGCTTGGCTTCAATGAGCAACACATTCGAAATGAAGATGAAGATGAAGAATTTAAATC 586
QY 602 TCCACCGGTTTTCATAGAT---TGATTAACATCCACATTTTAGATCAAAATAAGTTGAG 658
Db 587 TTTGCCAGTTTTCATAGATCCGTCATAAATATCCAGCAAAATTTGACTCCATTTATTG 646
QY 659 ACATACAAGGCTCGAAAGTTTATATAACCTTCGAGCGCTTGTGGTCCCACTGATAAAGT 718
Db 647 GAGAACCGGCTAAATTTCTTTTGCAATCTCGTCATCTCCCTATTACCGAGCGACCGCAT 706
QY 719 TCGGTCAACGTTTGTGTGACCGAATATATCAATCTACTAAAGAAAAGGTCCTTAAC 778
Db 707 TCGAACCCAGCTTCGTTTTCACCAATCCAAAATTAAGAAATTCGAAGGTTTGGATTGAGT 766
QY 779 CCAAGTCCCAACTTGGAGTACATCTCTTTTACGGTAACCTTGTGGTTATATATGAG 838
Db 767 CAGAGTTCCAAGTTTGTCTCATCTCTTTGTAGCGATTTGACGCTTATATGTGGC 826
QY 839 TTGCATAGCGAAATCACTTCGTAATAATAGGAGAAAGGCGAGACGAGTTAGAAC 898
Db 827 TGGCATAACGAATATTC-----ACAGCAGATGAAGACCAAGACGAGGATGCATT 880
QY 899 GTTCAATCAACGATGATGTGATCTCGTCTGTGATCCACCAATTTCCACAGCCCTACTT 958
Db 881 TTTCTGATTCGGTTCGATTAAGGCCACGATTAGATCCGCCGTTTCCCTGAAAATTAATT 940
QY 959 TGGTAACCTGTGGTGCACCATGTCTCCGACCTTAAAAAATGTCGTTTGTACTACGGAAA 1018
Db 941 CGGGAATGCTTATGCTACGCGTGCAGAGAAATGCGCGCGGAGAGCTGCTGGGAGAA 1000
QY 1019 TGGGTATGCACTTGGTGTAAAGTAAATTTGGAGAGTCTATATGCAAAATGATATATAAT 1078
Db 1001 AGGGTGTCTGCGACTCAGGTAATCGCGCGGAGATAAAGAAAGGATCAACGACAA 1060
QY 1079 GGAGGGAATCTTGAAGATGCCCGGAGATGGCATGA---ACCTTTTCATGATCCCGGTAG 1135
Db 1061 GAGAATATTAGAACCGGTGGAGAAATGGTCCCGGAGATTCGTAAGCGTTGCGAGAAATC 1120
QY 1136 GAAGATTGGTGTGCTGGTACACCTTAAGCTCAACTGTACGACTTTTGATTTTGGTGGG 1195
Db 1121 ATATTTTCCGGTGGAGATCGAGCAAGCTAGATCTTTACGGTGCAGATTTTGGATGGG 1180
QY 1196 GAAGCGCATAAAGTATGAGACTGTTTCAATAGA---CTAATAACGTCGATTTCTATATA 1252
Db 1181 GAAGCGGAGAAAGCAAGAAATATTGCGATTGATGGGAGAAATATGCATGACRCTTG 1240
QY 1253 TCGAAGCAAAACATCAGCAAGACTTTGAAATTTGGAATTTGAGTCTACCGAGTATGCAAT 1312
Db 1241 TAAAGCCAGGGAATTCGAAGGAGGATTTGAGGTTTGTCTTCTTGTGCTAAGGACAAAT 1300
QY 1313 GGAGCGTTTCTAGCATCTTTGATGAAGGATTAGAGATCAA 1355
Db 1301 GGATGCTTTTGTCTTATTTTTCACCTGGGAATTAATGGTTAA 1343

RESULT 6

E12754 LOCUS E12754 1622 bp DNA PAT 24-JUN-1998
DEFINITION Gentianatriflora mRNA for acyltransferase, complete cds.
ACCESSION E12754
VERSION E12754.1 GI:3251586

KEYWORDS JP 1997070290-A/2.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1622)
AUTHORS Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y.,
Yonekura,K., Mizutani,M. and Kusumi,T.
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 2 18-MAR-1997;
SUNTORY LTD
COMMENT OS Gentianatriflora
PN JP 1997070290-A/2
PD 18-MAR-1997
PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO,
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
C12R1:865),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1622
FT /organism='Gentiana triflora' FT
FT /variety='japenica'
FT /tissue_type='petal'
FT /clone='pGAT106'
FT CDS 35..1474
FT /product='acyltransferase'.
FEATURES
Source
I..1622
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 478 a 345 c 322 g 477 t
ORIGIN

Query Match 9.2%; Score 138.2; DB 10; Length 1622;
Best Local Similarity 48.6%; Pred. No. 2e-20;
Matches 676; Conservative 0; Mismatches 643; Indels 72; Gaps 8;

QY 1 TGAACATCTCGAACATGCGCGAATATCGCCCCCTCGGCACCATCGGCATCGCTCGT 60
Db 57 TCAAGTCTTGGAGAATGCGGTGCGCCACACCGGACGCGTCCGCGAGTTTACAG 116
QY 61 TATCTCTTACTTTCTCGACATTAATTTGGCTACTCTTCCCTCGGGTCCACCATCTTTTCT 120
Db 117 TCCACTGTCTTTTCGACATGCGATGGTTGATCTCTGATGAGAACACCATCTGCATT 176
QY 121 TCTATGACTTTCCACATCTTAAATCCCATTTTCATGGACACTATTTGTTCCAGGCTAAAC 180
Db 177 TCTACAGATTCGCGCATCTTGTCCCAACTCTAAATTTATCATTTTATCCCA---TTAAAT 233
QY 181 AATCTTTATCGTCACATCTTAAACATTTTTCGTTTCCGTTTGTAGTAATTTGATCTATTTC 240
Db 234 CGTCCCTTTCCCTTGTCTCAAAACACTTCTCTCCGTAGCCGGGAATTTGATTTGGCCGG 293
QY 241 CTAACTGATGGTTCGGGTTTAAATAAAAAACAGAAATAAAACAGCTTTGAAGGTGATT 300
Db 294 TAGATCTCCG-----ATAGATGCGGAGTTGCGTTACAAAGAGGGACT 341
QY 301 CTGTTGTGTTACTTTTGCAGAAATGTTGCTTCTGACTTTAATAATTTGACAGAAATCATC 360
Db 342 CCGTTTCTTTAACAATTCGAGAAATCGAGCATGGATTTGATTTATCTCGCGCGAGATCATC 401
QY 361 CTGAAATGTGAAAACCTTTTATCCACTTGTACCTTCATTTGGGAAATGCAATCAATAT 420
Db 402 AGAGGGATCTTTAAATTCACGATTTGATTCGCGAGCTGCCAGAACCGGATTTGTAACCT 461
QY 421 GTGATTCGTCACGGTCCCACTTTTTCACCTTCAAGTGACGCTTTTTCGCGGGTTCGGGTA 480

```
Db 462 CGGCGAGCAAGTATTACCACTTTTGGTTTACAGGTGACGGTGTCTCCACACCGGTA 521
Qy 481 TATCACTAGGAATACGAATCATATAGCTTGGTGACGCTAGCACGGGTTCACACTTTT 540
Db 522 TATGATTTGGAGCAATCTTATCAAGTCTTGGTGATGCCAGTCTTCTTTCGCAATTTA 581
Qy 541 TGAAGGTTGACCTTCGATTTATCAATCTGTTGTAGAT-----CGGT 582
Db 582 ATAAATTTATGGTCTTGGTTGACAAATCCAATGGAGATTCTATAAGTTCTCTCCACTTT 641
Qy 593 CTTTTTAAACGAAGATCTCCACCGGTTTTTGATAGATTGATTAACATCCACATTTAG 642
Db 642 CTTCTCTACCTATGACACAGATCTGGTGCAAGATCCATTTTATATTCGTCGCAAAA 701
Qy 643 ATGAAATTAAGTTGAGACATACAAGGCTCGAAAGTTTTTATAAACCTTCGAGCCTTTG 702
Db 702 TCTACAATGAAGAANAATGCTCAATCTCAGGGCACACCTACTGTTCTAAATCCAGCA 761
Qy 703 CTCCTACTGATAAAGTTCCGTCACGTTTGTGTTGACCCGAACCTAATATCAATCTACTAA 762
Db 762 TTTCTAAAGATGAAGTTCCGACCACTTCACTACACCTTATGATATCATGAAGCTCA 821
Qy 763 AGAAAAGGTTTAACCAAGTGCCAAACTT-----GGAGTACATGTCAT 807
Db 822 AGAAATTCATTTTCGTCAAAAAATCGCAACTTAACCGGTAGTAGTAATTAATCTGTCAA 881
Qy 808 CTTTTACGCTACTTGTGTTATATATGAGTTGCATAGCGAATCACTCGTAAATAAG 867
Db 882 CTTTCACCGGTGACATCTGCATGATCTGGAGATGCTTCTCGAAATTCATTAGACACGTCG 941
Qy 868 GAGAAAGAAAGGGCGAAGACG-----AGTTAGAACAGTTTCATATCACTACCATGATGTC 921
Db 942 TAAGAGAGAAGTGGAAGAGATAAACAATGACGACCAACTTATGCTTTTCATCACTGCC 1001
Qy 922 GATCTCGTTCATGCCCAATTTCCACAGCGCTACTTTGGTAACTGTTGGTGACCATGTG 981
Db 1002 GACAACGTTTTGCTCCGCGATACCTCAAAATTAATCTTTGGAATTTGCATAGTGCCTTGA 1061
Qy 982 TCCCGACCTTAAAAAATGCTGTTTGACTACGGAATAATGGGTATGCTACCTTGTGCTAAAG 1041
Db 1062 TGGTGGGATCGACTCATGACCAACTTTGAGGAATGAAGGGTTGTCGGTAGCTGCAACCG 1121
Qy 1042 TAATTTGAGAGTCTATATGCAAAATGATATATAAAGGACGGAATCTTTGAAGATGCCG 1101
Db 1122 CCATCGAGATGCTATCCATAAGAGGTTTACATGACTACGAAGGAATCTCGAGAGGAGAT 1181
Qy 1102 CGAGATGCGATGAACCTTTCATGATCCGG-----CTAGGAAGATTGGTGTGCTG 1152
Db 1182 GGATATCGGCGCGCGATCAACATCTCGCGCACCAAGGTCACGCTCATTTATGTGCTGTG 1241
Qy 1153 GTACACCTTAAGCTCAACTTGTACGACTTT-----GATTTTGGTGGGGAAGCGCATAA 1206
Db 1242 GATCGGACACGCAATGTGATGATTTTGGATGAGATTTTGGTGGGGAAGCTTGAAA 1301
Qy 1207 AGTATGAGACTGTTTCAATAGA---CTATAATAGCTGATTTCTATATAATCAAGCAAAA 1263
Db 1302 AGCATGATCTGTTTCAACTAATCTTCGGCAACACATAATTTTGATCTCTCGGTCAGAA 1361
Qy 1264 CATCAGCACAGATCTTGAANAATGATGATCTACCGAGATGCAAAATGGAGGCGTTTT 1323
Db 1362 GATTTAAAGGAGCACTTGAGCTTGCACTTCTTTGCTTGAAGATAGGATGACGCAATTTG 1421
Qy 1324 CTAGCATCTTT 1334
Db 1422 CCACCATTTTT 1432
```

RESULT 7

AB026494

LOCUS

DEFINITION Gentiana triflora GAT106 mRNA for acyltransferase homolog, complete cds.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AB026494
AB026494.1 GI:7415596
acyltransferase homolog.
Gentiana triflora cDNA to mRNA.
Gentiana triflora

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.

REFERENCE

AUTHORS

TITLE

Yonekura-Sakakibara, K., Tanaka, Y., Fukuchi-Mizutani, M.,
Fujiiwara, H., Fukui, Y., Toshihiko, A., Yamaguchi, M. and Kusumi, T.
Molecular cloning and biochemical characterization of
hydroxycinnamoyl-CoA:anthocyanin
3-O-glucoside-6-O-hydroxycinnamoyltransferase from Perilla
frutescens and diverse plant acyltransferase homologs

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Unpublished (1999)
2 (bases 1 to 1622)
Tanaka, Y. and Yonekura-Sakakibara, K.
Direct Submission
Submitted (21-APR-1999) to the DDBJ/EMBL/GenBank databases.
Yoshikazu Tanaka, Suntory Ltd., Institute for Fundamental Research;
Wakayama-dai 1-1-1, Shimamoto, Osaka 618-8503, Japan
(E-mail: toshikazu_tanaka@suntory.co.jp, Tel: 81-75-962-8807,
Fax: 81-75-962-8262)

FEATURES

source

1. .1622

/organism="Gentiana triflora"

/db_xref="taxon:55190"

35. .1474

/gene="GAT106"

35. .1474

/gene="GAT106"

/codon_start=1

/product="acyltransferase homolog"

/protein_id="BA93452.1"

/db_xref="GI:7415597"

/translation="MAGNSEDIKVKRCRVAPPDVAEFTVPLSFDMRWLISDAEH
HLHFYRPHCPNPKFTIISIKSLSLVLKHLPLAGNLINPVDSSDMPRLRYKKG
SVSLTIAESSMDFYLAGDHORSYKENDLIPLPEIPTVSGDEVLPFLALQVTFN
TCIGIRNLHQLVGLDASSFLHFNKLWLVKDSNGDSLKFLPLSLPMDRVSVDPEH
IRRIYNERKLLKSQGTPTVLNPAISKDEVRATFILPIDIMKLKFKISSKNRLTGS
SNYLSFTVTISALINWTCLSLDIVVREKVEDKHAANLCAFINCRORAPPQNY
FGNIVCEPMVGSQNRNVDHDFADFGWGKLEKHESVSTNPSTLILISRSRRFKGAL
APSTLIYVYVGSQNRNVDHDFADFGWGKLEKHESVSTNPSTLILISRSRRFKGAL
GISLPKNRMDAFATIFTNFINSIHVRPL"

BASE COUNT 478 a 345 c 322 g 477 t

ORIGIN

Query Match 9.2%; Score 138.2; DB 12; Length 1622;
Best Local Similarity 48.6%; Pred. No. 2e-20;
Matches 676; Conservative 0; Mismatches 643; Indels 72; Gaps 8;

Qy 1 TGAACATTCGACATGCCGGAATATCGGCCCTCGGGCACCACATCGCCATCGCTCGT 60
Db 57 TCAAAGTCTTGTAGAAATGCGGTGTGGCCACCACCGCGCGTCCGCGAGTTTACAG 116
Qy 61 TATCTCTACTTTCTCGACATTTACTTGGCTACTCTCCCTCCGCTCCACCATCTTTCT 120
Db 117 TCCCACTGCTGTTTTTGCACATGCGATGGTTGATCTCTGTATGACAGAACCATCTGCATT 176
Qy 121 TCTATGACCTTTCACATCTCTAAATCCCACTTTTCATGGACACATATTGTTCCCAAGGTAAAC 180
Db 177 TCTACAGATCCGCACTCTTGTCCCACTCTAAATTTATCATTTTCATCCA--TTAAT 233
Qy 181 AATCTTATCGGTCACTCTTCAACATTTTTTCCCGTTTGGCTAGTAATTTGATTTATTC 240
Db 234 CGTCCCTTTCCCTTGTCTCAACACTTTCTTCGTTAGCGGGAATTTGATTTGGCGG 293
Qy 241 CTACACTGATGGTTCGGGTTTTTAATAAAACACAGAATAAACAACGTTCAAGGTGATT 300
Db 294 TAGATTCCTCCG-----ATAGATGCCGGAGTTCCGTTACAAGAAAGGGACT 341

QY 301 CTGCTGTGTTACTTTTGCAGAAATGTTGCTTGACTTTTAATAATTTGACAGAAATCATC 360
Db 342 CCCTGCTTTTAAACAATTCAGAAATGCAGATGATGTTGATTATCTCGCGGAGATCATC 401
QY 361 CTCGAAATGCTGAAACACTTTTATCCACTTGTACCTTCATTGGGAAATGCAATCAATAT 420
Db 402 AGAGGATGCTTTATAAATTTCAAGATTTGATTCCGAGCTGCAGAACCGATTTGTAACT 461
QY 421 GTGATTCGCTACGGTCCCACTTTTCTCACTTCAAGTGACGTTTTTTCGGGCTCGGGTA 480
Db 462 CCGGCGCAGAGTATTACCACTTTTTCCTTTACAGGTGACGGTGTCTCCAACACCGGTA 521
QY 481 TATCACTAGGAATGAGGAATCATCATAGCCTTGGTGACGCTAGCAGCGGTTCACACTTTT 540
Db 522 TATGCAATGAGCAGCAATCTTTCATCAAGTTCTTGGTGATGCCAGTTCTTTCTGCAITTTA 581
QY 541 TGAAGGGTGAGCTTCGATTTCAATCTGGGTAGAT-----CGGT 582
Db 582 ATAAATTTATGGTTTTTGGTTGACAAATCCAATGGAGTTCAATTAAGTTTCTTTCCACTTT 641
QY 583 CTTTTTTAAGGAAAGGATCTCCACCGTTTTTTCATAGATTGATTAAACATCCCACTTTAG 642
Db 642 CTTCTCTACCTATGTAGCAGATCTGTGGTGCAAGATCCATTTCAATTCGTGCGAATAA 701
QY 643 ATGAAATTAAGTTGAGACATACAGGCTCGAAAGTTTTTATAACCTTCGAGCCTTGTG 702
Db 702 TCTACAATGAAGAAATGCTCAAAATCTCAGGCGCACACTACTGTTCTAAATCCAGCA 761
QY 703 GTCCCACTGATAAGTTTCGGTCAAGCTTTGTTGACCGGAAGTAAATCAATCTACTAA 762
Db 762 TTTCTAAGATGAAGTTTCGAGCAGCTTCTCTACACCTATTGATATCATGAAGCTCA 821
QY 763 AGAAAAAGGCTTTAAACCAAGTCCCAAACTT-----GGAGTACATGTCAT 807
Db 822 AGAAATTCATTTTCGTCAAAAATTCGCAACTTAACCGGTAGTAGTAATTAATCTGTCAA 881
QY 808 CTTTTACGGTAACCTTGTTGTTATATGAGTTGTCATAGCGAAATCACTCGTAAAAATAG 867
Db 882 CTTTTCAGGTGACATCTGCACCTGATCTGGACATGCTTGTGAAATCATTAGACACCGTCG 941
QY 868 GAGAAAGAAAGGCGGAAGCG-----AGTTAGAACAGTTTCAATACCACTGATTGTC 921
Db 942 TAAGAGAGAGGTGGAAGAGGATAAACATGCAGCAAACTTATGTGCTTTTCATCAACTGCC 1001
QY 922 GATCTGCTCTGTATCCACCAATTCACACGCTTACTTGTGTAAGTGTGGTGACCACTGTG 981
Db 1002 GACAACTTTTGTCTCGCCGATACCTCAAAATTAATTTTGGAAATTCATAGTGCCTTGA 1061
QY 982 TCCCGACCTTAAAAAATGTGCTTTTGACTACGGAAATGGTATGCACCTTGGTGCTAAAG 1041
Db 1062 TGCTGGATCGACTCATGAGCAACTTGTAGGAAATGAGGGTTCTCGGTAGCTGCAACCG 1121
QY 1042 TAATTGAGAGTCTATATGCAAAATGATATATAATAGGACGGAATCTTTGAAAGATGCCG 1101
Db 1122 CCATCGAGATGCTATCCATAAGAGGTTCATCACTACGAAGGAATTCAGAGGAGATT 1181
QY 1102 CGAGATGGCATGAACCTTTTCATATCCCG-----CTAGAAGATTGGTTGCTG 1152
Db 1182 GGATATCGCCGCGCGATCAACATCTCGCGCACCAAGGTGCAGCGCTCATTTATGCTGTG 1241
QY 1153 GTACACTTAAGCTCAACTTCTAGCACTTT-----GATTTGGTGGGGAAGCGGATAA 1206
Db 1242 GATCCGACACCAAGCAATGTGCAATGATTTGATGCAGATTTTGGTGGGAAAGCTTGAAA 1301
QY 1207 AGTAGAGACTGTTTCAATAGA---CTATAATACGCGATTCTTCTATAAATGCAAGCAAAA 1263
Db 1302 AGCATGAATCTGTTCAACTAATCTCTCGGCAACACTAAATTTGATCTCTCGGTCCAGAA 1361
QY 1264 CATCAGCACAGATCTTGAATTTGGATTGAGTCTACCGAGTAGCAAAATGGAGCGTTTT 1323
Db 1362 GATTTAAAGGAGCACTTGAGCTTGGCATTTCTTGGCTAAGAATAGGATGGAGCGCATTTG 1421
QY 1324 CTAGCATCTTT 1334

Db 1422 CCACCATTTT 1432

RESULT 8
AC002560
LOCUS
DEFINITION
AC002560 101176 bp DNA PLN 19-JUL-2000
Genomic sequence for Arabidopsis thaliana BAC F21B7 from chromosome 1, complete sequence.
AC002560
AC002560.2 GI:9211167
HTG.
SOURCE
Arabidopsis thaliana
ORGANISM
thale cress.
REFERENCE
AUTHORS
Khan, S., Brooks, S., Buehler, E., Chao, Q., Johnson-Hopson, C., Kim, C., Shinn, P., Altafi, H., Bei, Q., Chin, C., Chlou, J., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N., Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukharsky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M., Yu, G., Federspiel, N.A., Theologis, A. and Ecker, J.R.
Genomic sequence for Arabidopsis thaliana BAC F21B7 from chromosome 1
TITLE
Unpublished
2 (bases 1 to 101176)
Ecker, J.R.
JOURNAL
Direct Submission
AUTHORS
Submitted (25-SEP-1997) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
TITLE
3 (bases 1 to 101176)
Ecker, J.R.
JOURNAL
Direct Submission
AUTHORS
Submitted (13-NOV-1997) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
TITLE
5 (bases 1 to 101176)
Ecker, J.R.
JOURNAL
Direct Submission
AUTHORS
Submitted (26-JAN-1998) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
TITLE
5 (bases 1 to 101176)
Cheuk, R., Shinn, P., Brooks, S., Buehler, E., Chao, Q., Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bei, B., Chin, C., Chlou, J., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N., Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, J., Liu, S., Mukharsky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M., Yu, G., Davis, R., Federspiel, N., Theologis, A. and Ecker, J.
Direct Submission
TITLE
Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA
REFERENCE
AUTHORS
Ecker, J.R.
JOURNAL
Direct Submission
AUTHORS
Submitted (15-JUL-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
TITLE
7 (bases 1 to 101176)
Cheuk, R., Shinn, P., Brooks, S., Buehler, E., Chao, Q., Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bei, B., Chin, C., Chlou, J., Choi, E., Conn, L., Conway, A., Gonzales, A., Hansen, N., Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, J., Liu, S., Mukharsky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H., Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M., Yu, G., Davis, R., Federspiel, N., Theologis, A. and Ecker, J.
Direct Submission
TITLE

QY 123 TATGACTTTCCACATTTCTAAATCCCATTTTCATGGACACTATTGTTCCCGAGCTTAAACAA 182
Db 37391 TTCTCTTTACCAAACTCAACTGAATCTTTCCCTCCCAAGACTTTGTACCCCAACCTCAACAT 37450
QY 183 TCTTTATCGGTGACCTCTTCAACATTTTTCCTCGGTTTTCGTTAGTAATTTGATTTGATTTCT 242
Db 37451 TCCCTCTCCATCACTCTCCCAACATTTCTTCCCTTACGCGGGTAAATGATTTATCCCGCT 37510
QY 243 AACACTGATGGTTCCGGTTTAAATAAAACAGAAATAAACACGTTTGAAGGTGATTC 302
Db 37511 CGTCCGACCTCCATATTTACACTACAACGCTG---GGAAGACTCTTTGTTTTCACC 37567
QY 303 GTTGTGGTTTACATTTTCAGAAATTTGTTGTTGACTTTTAATAATTTGACAGGAATCATCT 362
Db 37568 GTAGCAGAGCTTACTGAAACCGACTTCGACCACTGAAGTCTGATTTCCCTTAAAGATATC 37627
QY 363 CGAAATGTGAACACTTTTATCCACTTGTACCTTTCATTTGGGAATGCAATCAATATGCT 422
Db 37628 AGTGTGTTGCATGGGGTCTTGGCCAAAGTTTACCTCTCCTCCTCAGCTCTCTCCGGAAGGAAT 37687
QY 423 GATTCGGTCCACGCTCCCACTTTTTCACCTTCAAGTGACGCTTTTTCGGGGCTCGGTATA 482
Db 37688 CA-----CATGGACCAATATGGAATGCAAGTCAACCATCTTCCCTGGAGCTGGAATC 37741
QY 483 TCACTAGGAATGACGAATCATATAGCCCTTGGTACGCTAGCAGCGGTTCAACTTTTTCG 542
Db 37742 TGTATAGGCACTCAGCTACACATGTTGTAGCAGATGGAGTCACTTCACTCATTTTCATG 37801
QY 543 AAAGGTGGACTTCGATTAATCAATCTGTTGATGATGCTTTTTCATGAAAGGATCT 602
Db 37802 AAGTATTGATGTCAATGACCAATCCAGCGCAAGATCCGCGCCACGGTTCTTCTTACCC 37861
QY 603 CCACCGGTTTTGATAGATTGATTAACATCC-----CACATTTAGATGAAT 650
Db 37862 TCTCTGCCATTTCACAGCTGCGAACAATATCAAGGACCCAGGAGAGTAGCCCGCAGGA 37921
QY 651 AAGTTGAGACATACAAGGCTCGAAAGTCTTAAACCTTCGAGCCTTGTGGTCCCACT 710
Db 37922 CATTTGGACGGTTTGGAGCCCAAACTCTGCNAACACAGTTTCAATGTTAGCCCTGAG 37981
QY 711 GATAAAGTTCCGTTCAACGTTTGTGTTGACCCGAACTTAATCAATCTATAAGAAAGAA 770
Db 37982 AACATGGTCAGAGCCCACTTTTACATTTAGCCGGAAGCAGATAGACAATCTCAAGAGTTGG 38041
QY 771 GTCCTTAACCAAGTGCACAACTTGGAGTACATCTCATCTTTTACGCTACTTGTGGTTAT 830
Db 38042 GTTACAGACGCTGTGAGATCAATCTCTGTTTCTTACCTTCTGTTGAGCTTAGCTTTC 38101
QY 831 ATATGGAGTTGCATAGCGAAATCACTCGTAAATAATAGGAGAAAGGCGGCAAGACGAG 890
Db 38102 ATTTGGGTAAGCTTGATTAGACACTTGTTCAGACAGTGAACAGAGGCCAAGGACGAA 38161
QY 891 TTAGAAGCTTCATATCAATCACTACATTTGATTCGATCTGCTTGTAT-----CCACCAATT 944
Db 38162 GTCTTTCACTTGATGATCAACCTGGATTTCGGAATCGTCTCAAGTACACACCAACCAATA 38221
QY 945 CCCACAGCTTACTTTGGTAACCTGCTGACCATGTTCTCCGACCTTAAATAATGCTGTT 1004
Db 38222 CCACAACATATCTTTGGCAACTGTATGGCTCTCTGTTATCGTATCAGTCAAGAAACAGAT 38281
QY 1005 TTGACTACGGAATAATGGGTATGCATTTGTTGCTTAAAGTAATTTGGAGTCTTATATCAAA 1064
Db 38282 TTGCTAGGAGAAAATGCGCTCATGCTGCTTTCAGATGCAATCACCGAGATCAAGAT 38341
QY 1065 ATGATATATAAAGGACCGGAATCTTTGAAGATGCGCGAGATGGCATGAACCTTTTCATG 1124
Db 38342 ATGTTATCAAGCATCTGTTGAAGACAGCACCAAGATGGGACAGGAGTACGTAATGCG 38401
QY 1125 ATCCCGGCTAGGAG---ATTGGTGTGCTGGTACACCTTAAGCTCAACTTGTACGACTTT 1181
Db 38402 GTCATGTCCTCATATCCAACTCAATTTGAGGAGCTCCGAAATTTGGGACTGTATGACATG 38461
QY 1182 GATTTTGGGTGGGGAAGGCAATTAAGTATGAGACTGTTTCAATAGACTATAATACGTCG 1241

Db 38462 GATTTTGGATTAGGAAAGCCCTGCAAAATGGAGATAGTCACATAGAAAAGGTGTTCT 38521
QY 1242 ATTTCTATAATGCAAGCAAAACATCAGCACAAAGATCTTTGAAATTTGATTTGATCTACCG 1301
Db 38522 ATCGCATTTCTAGAGTTCAGAGACGCGCAATGGAGTTTGAGATTTGGAATAGCACTAGAG 38581
QY 1302 AGTATGCAATGGAGCGGTTT 1322
Db 38582 AAGAAGAAATGATGTTT 38602
AC003027 119914 bp DNA PLN 30-JAN-1999
Arabidopsis thaliana chromosome I BAC F21M11 genomic sequence,
complete sequence.
AC003027 AC003027.1 GI:4079614
HTG.
SOURCE
ORGANISM
Arabidopsis thaliana
thale cress.
REFERENCE
AUTHORS
Fiederspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altati, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vysotskaia, V.,
Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
Unpublished
JOURNAL
REFERENCE
AUTHORS
Fiederspiel, N.A., Palm, C.J., Conway, A.B., Kurtz, D.B., Conway, A.R.,
Au, M., Araujo, R., Buehler, E., Dewar, K., Feng, J., Kim, C., Li, Y.,
Oji, O., Osborne, B., Shinn, P., Sun, H., Toriumi, M., Vysotskaia, V.,
Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
Direct Submission
TITLE
JOURNAL
Submitted (22-OCT-1997) Biochemistry, Stanford University/DNA
Sequencing and Technology Center, 855 California Avenue, Palo Alto,
CA 94304, USA
3 (bases 1 to 119914)
Fiederspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altati, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vysotskaia, V.,
Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
Direct Submission
TITLE
JOURNAL
Submitted (30-DEC-1998) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
4 (bases 1 to 119914)
Fiederspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altati, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vysotskaia, V.,
Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
Direct Submission
TITLE
JOURNAL
Submitted (30-JAN-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
COMMENT
On Dec 30, 1998 this sequence version replaced gi:2734094.
Bases 1-9462 of clone F21M11 overlap with bases 68998-78259 of
'TAMU' BAC clone F20D22 (AC002411) and bases 119525-119914 of clone
F21M11 overlap with bases 1-389 of 'TAMU' BAC clone F21B7
(AC002560).
e-mail for correspondence: arab@sequence.stanford.edu
Genes with similarity to proteins in the databases are described as
'putative', 'like' or 'similar to'. Genes that have EST
similarity but no significant protein similarity are described as
'unknown proteins'. Genes that are annotated based only on gene
prediction software are described as 'hypothetical proteins'.
The software programs used to predict genes include: Grail

(Informatics Group, Oak Ridge National Laboratory,
<http://compbio.ornl.gov/section/index.html>), GENSCAN (Chris Burge,
<http://genome.stanford.edu/~chris/GENSCANW.html>), Fexa (V.Solovyev
 & A.Salamov, Sanger Centre, <http://genome.sanger.ac.uk/>), and
 NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, <http://www.cbs.dtu.dk/NetPlantGene.html>).

FEATURES

source
 1. 119914
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="F21M11"
 1. 7000
 /note="Proteins in this region are annotated in the F21M11
 entry, AC002411."
 misc_feature
 1. 9262
 /note="Overlap with bases 68998-78259 of 'IGF' BAC clone
 F21M11."
 gene
 7684..10913
 /gene="F21M11.1"
 join(7684..7866,7965..8027,8122..8160,8250..8381,
 8458..8728,8875..9893,10003..10256,10344..10410,
 10501..10551,10685..10781,10885..10913)
 /gene="F21M11.1"
 /note="Hypothetical protein"
 /codon_start=1
 /protein_id="AADI0665.1"
 /db_xref="GI:4204284"
 /translation="MAPNLRIKKACDAMKLLGSETKTRAFPLRKLKTYENNNDFTIEE
 DAYKLLDALFDADAQSTBKNKEEKKKKSRVATSGRRKAPPEPLVQDEE
 DMDDEFFLRDLRRGRASSSSSSSSNNEDLQTPDEEDDQGVTELPPLKRY
 VRRNGEAPLVYNNAPSSSRLSMEPEPVMVLLPAHPETKVSASALVILND
 EPNIDHKPVTSDGNCAPMLEGKSNHVEQDWEKIDILNDDTAMDVSPSALICES
 SEHKVAASVELASSTGEAKICLSFAPAGETNHLPSMEDLRAMEKCLSKYSKI
 VHPFSLVGLFMKMCSDITDAGENEVEIPWNEINEKVPFRYPHSHFVQDAPVIFS
 ECETISGDQKAISSKMDITAGEENVEIPWNEINEKVPFRYPHSHFVQDAPVIFS
 LSSFDSQSCSTICEICLASCSNCAIGVDFNGFAYTLDLGLLKEEPLARISEARDQ
 RKQVLECECPLEAKKVEILPCKGLHGAIKCEKFCGCTKRCGNRVQVGMHN
 KLVQFFTPNKGWGLRTLEKLPKGAFCIEYIGELTIPELYQSFEDKPTLPVILDAH
 WGEERLEGDKALCLDGMFYGNISRFNLHRLCLDANLIEIPQVETPDQHYHLAFTT
 RDIEMAEALWDGIDRDNDSMLKPFDCIGSRFCNRKTRKTMQILNKA"
 join(13033..13488,13717..13926,14139..14514)
 /gene="F21M11.2"
 13033..14514
 /gene="F21M11.2"
 join(13061..13488,13717..13926,14139..14316)
 /gene="F21M11.2"
 /note="Similar to acid phosphatase; Location of ESTs
 110C2T7 , gb|T42036, and 110C2XP, gb|A1100245"
 /codon_start=1
 /protein_id="AADI0666.1"
 /db_xref="GI:4204285"
 /translation="MDRTMFLSTIALLGVVVSAGDWNILNOLRGIGSSSSONGIVS
 KGTLDLKGCESNRINVEHNIRKEDVPQECVSHIKOYMTSQYKDDVARTVDEVI
 LHFSGMCCSKCDGMDAIFDIDTLSTIPYHKKNCFEGEKLNSKFEDWTQKKI
 APAPVHKLVYHDIRERGIKIFLISSRKEYLSATVDNLQAGYVGSNLMVLAGLEQ
 QKEVQKVKSEKRWLMSLGYRVWVGMDGDNSSAGCPLPRTPLKLSNIYYVA"
 complement(join(14004..14312,14565..14641,14787..15831,
 15906..16006,16327..16362))
 /gene="F21M11.3"
 complement(14004..16362)
 /gene="F21M11.3"
 complement(join(14567..14641,14787..15831,15906..16006,
 16327..16362))
 /gene="F21M11.3"
 /note="Unknown protein; Location of ESTs 40C3TT,
 gb|AA728590 and 40C3T7, gb|T04573"
 /codon_start=1
 /protein_id="AADI0667.1"
 /db_xref="GI:4204286"
 /translation="MLDQISGNHTHEKLSVETAEPHLNDNRVHIVEIPKASVITPE
 ICDEAEKCSPTISRKRVTFTDSKVITYEHVSVSEESVELSEKNEEVEKRSKLKSK

gene
 CDS
 complement(join(16835..20238)
 /gene="F21M11.4"
 complement(join(16835..17185,17274..17392,17491..17656,
 17795..17885,17982..18079,18175..18361,18504..18604,
 18705..19049,19134..19349,19439..19791,19862..19967,
 20188..20238))
 /gene="F21M11.4"
 /note="Hypothetical protein"
 /codon_start=1
 /protein_id="AADI0669.1"
 /db_xref="GI:4204288"
 /translation="MGSSSPPEARARAOVPSMILIFLEIITVHYVYNNRRLNDRDLVLSA
 NLNIPKVRTRDLRFEMSVISYKSLNAAVSVHLPOIQIPNDCNYKNDALNNNSPKH
 GESEDSMTDKVSKRGSDSSRGLSPLETSSEDPKRGHQTDEKQLSHLLLYE
 FESEYDAANHTPESYTDQAANKNQMDYFASQPSNAARKRICGDSFIOESSPNPKTOPD
 TLLRLMESLRSDOPTDYVKAQNHQMVYFSEWVLYSFVLCIERSNNLRGCVIFCQOOLPK
 SHTEQDSKRRDITASDAMENHLKVPKRENNLMQKSDADICNCKCSANSDQDSEKIS
 KALEQTSNNITICGQCSARVSEATGEMLHYSRGPVDDGDIFRSNVIHVHSAIEWA
 WQVEYEGDTYKLNKLAELARGMKIKCTKSLKGAALGCFVKSCRRSLEQTPAFTKELVLCGR
 SALSKSDKLMESLAVRFENATISRYWNPSTVTHVIASTDEKGACTRTLLKVLMLGKNGW
 IINNAMKSLKASQVDEEFPFIQIDTQCQDGPRTARLAETNKKPLEGGLKFFYF
 GDYPTKTELDQNLVYAGVTILNDELGAESSNNVNDQRSSIVVYINIDPPHGICAL
 GERVTIWRANDAEALASQTSRLVGTWVLESIAGYKLHPVIG"
 20752..24647
 /gene="F21M11.5"
 join(20752..20994,21362..21497,21596..21740,21825..21994,
 22102..22178,22542..22758,22920..23064,23343..23442,
 23599..23693,24021..24072,24227..24298,24441..24647)
 /gene="F21M11.5"
 /note="Hypothetical protein"
 /codon_start=1
 /protein_id="AADI0668.1"
 /db_xref="GI:4204287"
 /translation="MGANSKSVTASFTVIAVFFLICGRTAVDETEFEHGDYSKLSGI
 IIPFASTQURAWSILDCPYTDLDFNPLDVLWDTTKLLSAVNCWFCMVLDDPYNQTD
 HPECKSRDPSGLSAITELDPGYITGPPLSTVMKWLKWCVEFGEANAIVAPVDMRLS
 PTKLEERDLYFHKLKLTFTALRGGPSIVFAHSMGNVRYFLFWLRLEIAPKHYL
 KWLDOHIAVFAGAPLLGSVEAIKTSVGTGTLGVPSEGTARLLSFSASSLWLMFP
 SKNCKGNTFTWTFHSGCAAKVDEARVHCDDEEYQSKYSGMPTNINIEIPDSARELA
 DGLFLKAIEDYDSDSKRMLHQLKYYPPFVIRNIAHRSSLAGLLEHDDVFNPLTPW
 ERPIKRVCIYGAHLKTEVGYIFAPSGRPYDNNWIIITIIYEGLSLSVRSQVVDG
 NAGPITGDETPYVYSLSWCKNLWLGPKVNIITMAPOILIGIKIQOQPEHGDGSDVHVELNVD
 HEHGSDIANNTRKAPRVKITYFVDESISIPGKRTAVWELDKSGY"
 27777..28734
 /gene="F21M11.6"
 27777..28734
 /gene="F21M11.6"
 28007..28465
 /gene="F21M11.6"
 /note="Unknown protein; Location of ESTs 20312T7,
 gb|H76794 and 203124XP, gb|AA605510"
 /codon_start=1
 /protein_id="AADI0670.1"
 /db_xref="GI:4204289"
 /translation="MNTKTMRLPPRRVLTADKKRERDAFISSTVDNPPETAKFPSPPP
 KLVPPVNPISKSSATAAPIGSNOLMAGYLSHEVLTOTGLFGEAGNWARAQAES
 KIKPSHTVEPAEECEPKRKRYREVANLLRSDGALPGIVNPAQLARFLKL"
 complement(29264..32033)
 /gene="F21M11.7"
 complement(join(29264..31015,31312..31414,31484..31587,
 31782..32033))

Query Match 5.88; Score 88.2; DB 12; Length 119914;
 Best Local Similarity 45.08; Pred. No. 1.9e-09;
 Matches 621; Conservative 0; Mismatches 718; Indels 42; Gaps 6;

Db	1002	CAAGCCAAACTGAAGCTGAAGCTGCAGCAGGTAATGATGAATTAATAATATATTGT	1061
Qy	902	CATAATCACCATTGATTCGATCTCGTCTTGATCCACCAATTCACACAGCCTACTTTGG	961
Db	1062	GGGATTCACTCGGATATAGGACCGTTGGATCCTCCATTCACATTAATTTTGG	1121
Qy	962	TAACGTGGTGACCATGTGCCCGACCTTAAANAATGCTGTTTGATACGACGAATGG	1021
Db	1122	TAACGTGAATGGAGACATTTGTGAGACTGCAAGAGCAAGTGATTTCTGTTCAAGAAATGG	1181
Qy	1022	GTATGCACCTGGTCTAAAGTAATTTGGAGA---GTCATATGCAAAATGATATAATAA	1078
Db	1182	GTTGCTCTTTGTCGAGAGATGTTAAAGTGATATGCTCAAGGGATCGATGCGGATGCCAT	1241
Qy	1079	GGACGAATCTTGAAGATGCGCGAGATGCGATGAACCTTTTCATGATCCCGCTAGGAA	1138
Db	1242	TGAAGCCATGATGATAGGTTTCAGAAATATGGAATTTCTGAAGAAGAGCAGCATAT	1301
Qy	1139	GATTGGTGTGCTGACACCTAAGCTCAAGCTTGTACGACTTTTGATTTGGGTGGGGAA	1198
Db	1302	TTTTTCTGCTGCTCGACCAATTTGATGTTTACGGTTCGGATTTTCGGTGGGGCAG	1361
Qy	1199	GCGCATAAGTATGAGACTGTTCAATAGACTATTAATACGTCGATTTCTATAAATGCAAG	1258
Db	1362	CCCCAAGAGTGGAGATGTTGTTCAATAGATAGGACACAAGCCATCTCTTTGGCAGAGAG	1421
Qy	1259	CAAAACATCAGCACAAAGATCTTGAATTTGGATTGAGTCTACCGAGTATGCAAAATGGAGCG	1318
Db	1422	AAGAGATGAGGAGCGCGCTTGAGTTGGAGTTGTTTATAGAGAACAACAATTTGAGGT	1481
Qy	1319	GTTTCTAGCATCTTTGATGAAGGATTAGAGAGTCAAGCTTTTCATTTGATGATCAT	1372
Db	1482	TTTTGAATCTGTATTTGCTGAGTCAAGAAATGATCTCTGTTTAATTAATGAT	1535
RESULT 11			
AB016892	66237 bp	DNA	PLN
LOCUS	AB016892	Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone:MXF12.	27-DEC-2000
DEFINITION	AB016892	BA00015	
ACCESSION	AB016892.1	GI:3449333	
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			

NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://gremlinl.zool.tastate.edu/cgi-bin/sp.cgi) genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K1566 and the 3' clone is K3K3.	
FEATURES	
source	location/Qualifiers
	1..66237
	/organism="Arabidopsis thaliana"
	/strain="Columbia"
	/db_xref="taxon:3702"
	/chromosome="5"
	/clone_lib="Mitsui p1"
	2613..5234
	/note="gene_id:MXF12.2"
	/codon_start=1
	/evidence=not_experimental
	/product="receptor-like protein kinase"
	/protein_id="BAB10824.1"
	/db_xref="GI:10177545"
	/translation="MIRHALLIFSLVSTPIVGBGATSTVEPTDVLFLNCGDTSNNVD VSGRNTAENQKILSSNLVNASFTAQAOYSGVSOIPYMTARIFRSEFTYSPVTPG SNELRLFYPTRYGSPQNAVKSFFSVKNGFTLLNFSADLTVKASKPQTEFIKEFI IPVTGLNLTFTPSLDSLAFAVNGIEIVSNRFSYKGGFDVITNNGSSVDFHLENST AFETVRLNKGKTVGDSGFRMRWSDDEILSESSGISPIVDPIKNTYKTEKPSYVA PDDVIATSRGMADHPQNLNMLTWLFTVDAGFSYLVRHFCETLSEVNKEQGVF SFIENOTATLEMDVFRMSGWIPMYLDYTVIAGSGRRHDLRLDHLPLVSNPKYF DYALLNGVEILKMNPDGNLGNPNPDLVSPDLIPNRTPIRKNKSHILPITLAVVG SLVLANFVGVVLVIMKKKKKSPSTNSNCPLPHGTDSTNTPAKSLPADLCRRFSI FEKSTATNDFDKLIIGVGFGSVYKQIDGGATLVAKREITLISNAGKEFEFESEM LSKLRHLVLSLIGYCEDENMLVLEYMPHGTLDKHLFRDRTSDPLSKRRLETC IGARGQLVHTGAKYTIHRDITKTNILLDENFVTVSDFGLSRVGTSASQTHVST VNGTCFGLDPEYRROVLTEKSDVSGFVGVVLEVLCCRPIMQSVPEQADLIRWVK SVNRCGTVDQIIDSLSADJTSLEKFECEIAVRCVQDGRMERPPMNDVYVWALEFALQ LHETAKKNDNVESLDLMPGSEVGTGTTDGEDDLFRTTGHVGKSTTTDDSVLVVVGDER SGSSWGVFSEINEPKAR"
	6701..7210
	/note="gene_id:MXF12.3
	unknown protein"
	/codon_start=1
	/evidence=not_experimental
	/protein_id="BAB10825.1"
	/db_xref="GI:10177546"
	/translation="MINLEKEAKEKQILIRDKRKIVHHVNTTLLVXIYBASAIIDRRH LFLKPYGPPNYVFIKAWLDDTKGYTVVVKDGFKNOPPLKIPLNFPQOYLVYVELLK GMSYRDPCTNRTVVMGRAKIRLSPLMTSYKERPISGIDLVGLNSDQCVVKKGYLYI AMKLHRSAD"
	7670..10111
	/note="gene_id:MXF12.4"
	/codon_start=1
	/evidence=not_experimental
	/product="receptor protein kinase-like protein"
	/protein_id="BAB10826.1"
	/db_xref="GI:10177547"
	/translation="MNCNVLFLSVLSVGTAGVTAAYHPTDVLFLNCGDTSNNVDSNG RNTVTSRQILSSNLVNASFTSEAKGVASRIPYMKARIFRSEFTYSPVTPGSGIF LQYLFYPTQYKSGFDVAVNSFVSXVNGFTLLRNFNADSTVQASITPLNSLTKKEFIIPV HRTNLNFTPTSKNLLAFVNGIEIVSMRPFYKGGFDVLRNNSDSDVDFDINSTAFE SVHRLNNGGOIVNEVDSDGMRBRLSDDSFGNSCSI VNPVGVKINYTEKPAVAVPD VYATSRMGNSNLMNLTGCMFLITVDAGYNLVRHLHFCETLPQVTKAGQRFSLFVED VMKATSDVIRLGGPPIPMYLDVSVYVFGESMIQDELRLDVLKPDQVOTYTDAIL KYATKILNDSDGNLARNPNELLVSDTSDPDSNVNTPIKRPHVLVILIVVGSVGGF SGVEILKNDSDGNLARNPNELLVSDTSDPDSNVNTPIKRPHVLVILIVVGSVGGF LATFIVTLMILLQMRKKKNKNSITMFKLLKQYIYAEKLTITKTSFSTVKGKGGF TVYRGNLSNGRTVAVKVLKDKNGDDFI NEVTSMSOTSHVNI VSLGFCYEGSKRAI ISEFHEGSLDQFISRNKSLTPNVTLYLGALGTAGCLEVLYHCKTRIVHFDIKPON ILDDNCPKPVADFGALKEKRESILSLIDRTGTIGYIAPEVSRMYGSGISHKSDVI SYGMLVDMIGARKNKVETTTCPDWIYKDLNGDGTOTWIGDEINEEDNKIVK KMLIVLSMCIIRPCPSDRPPMKNVMEGSLDALELPKPPKPSRHIESTLVESSLSDG

```
CDS
QEAERQQTQTLDSITII"
10819..13239
/note="gene_id:MXF12.5"
/codon_start=1
/evidence=not_experimental
/product="receptor protein kinase-like protein"
/protein_id="BAB10827.1"
/db_xref="GI:10177548"
/translation="MIFCLFVSEFLVSATAPYKPDVDFLINGGETDVPDNGHRT
WTOPEKNTLPKNSDNASFSSVSVYKEESGIQVPYMTARIPRSDTYSFPVSPGHKFL
RLFYPTSYKSGFDVAVNSFVYNDFTLLQNFSDADLVAKASIPESKSLIEFIVPVY
TLNLNLTFRPSNLSFANGIEIVSMDFRYSKGGDDDLITNGVSLDEIDNSTASET
VHLNVGMDGDEVDNSGFMFRWLSDDEFLIGGVSPYMPDYNISYTDTEFTAYAPAY
VYSTRCMGNAQDTLNLNFTLWLTVDAGFSYLVRLHFEKYLKANKQRVFSIFLG
NOMAREMDRIQLSGGPRIPIYLDPRIVVGSSESGPRDLRLDLHLPLNDPEYFIAL
NGVELLKLNSGNLAIODNELKPNPLSSNLTPNHVTQOIKGKSHLLLVKLFIAVGP
GTGLATFVVLMLNRQMKRNRKEERVMEFKLLNMTYIAELKIKTSFSTYICGGG
FGTVVYGNSGRKVAVLKDLKQSAEDFINAVASQTSHTVNIIVSLLGFEQSKR
AIYVEFLNGSLDQPMRSNKSUTQDVTLYLGIALGARGLEYLHYGCKRIVHFEDIK
QNILDGLCPKVSDFGLAKICEKRESVLSMDTRGTIGYIAPEVFSRMVGRVSHKSD
VYSFQMLVIDMIGARKIEIVETVDSAASTVFPDWIYKDLDEGDEOTWIFGDEITKEEK
EIAKMIUVGLWCIOPCPSDRPSMRNRVMEHSGSLDALEIIPKPSMHISTEVIETESS
LSDGGEV"
Join(16412..16498,16859..16924,17003..17038,17085..17192,
17337..17461,17541..17640,17803..17974,18073..18191,
18285..18395,18488..18632,18712..18844,18934..19009,
19232..19297,19385..19492,19591..19698,19873..20070,
20164..20373)
/note="gene_id:MXF12.6"
/codon_start=1
/evidence=not_experimental
/product="ABC transporter-like protein"
/protein_id="BAB10828.1"
/db_xref="GI:10177549"
/translation="MGNKLLTGGSSKTHGSSSYRDPDLLQNEQDKPKANGSEGLND
LEHGVVEANQVGFGRFALICQYSIMVSOAKPDAGKIVIGTIALIGSTNLLVPKF
GMMIDVSRVKTPEQTESLIAVNAVVIILLVIGSICITALLRAWLFNSASRVP
ARLRKDLFRHLHQHIAFDVDTKCELLSRSEDFQIIKNAATNLSLSALRNTALI
GVGFNTSLMLTLLAVVVPVSVAVKQFGRYRELSHTTQAAAVALSAIRSEFGA
VRTVRSFAKESVMVSOYSKQVDETLKGLKQALVGLFEGGLNAFTLVITVSVYGA
LTVIGSMTVGAULTLIYSLTVGSSVSSLSLYTTAMKAAGSRVQFILDVRSMS
SGDKCPVNGPDGVDELNDVWFAYPSRPSMILKIGISLRLTPGSKVALVPGSGGKTF
TANLERFYDLPLKRIANGIEFLSMELSHOYLHKQISVISOEPILFNCSVEENIAYGFD
GEASPTDENAKMANAHEFIEAFPDKYNTVYGERGLRSGGQKQRIAIARALLNPS
VLLDEATSDALDAESEYLVODAMDSLMAGRTVLVIAHRLSTVTKTADCVAVISDGEVAE
KGTDELLSLNGIYNLKVROLQSSSVTLL"
Join(25348..26359,26401..26756)
/note="gene_id:MXF12.7"
/codon_start=1
/evidence=not_experimental
/product="anthocyanin acyltransferase-like protein"
/protein_id="BAB10829.1"
/db_xref="GI:10177550"
/translation="MVNEMESSLKVIDVARVTPSNSDSSSLTLPFLFFDLLWYKLH
AVERYIFKLTDAAPFFDSVTPNLKTSLSLSLSHSLYPLAGKLWPELDLPKPKIVYT
PNDVAFYVAESNADFSLRTGKEPFTTELYPLVPELHVSDSASAVSFQVTLFNPQG
FCISVNAHVAVLGKTTNFKLSWARTCKNODSELPQDLIPVYDRTVLKDPMDLDTKI
LNAHRAKVFTEGKEPENPKSLKLLWSPGDPFVRTYLNLTREDIQKRLERLAKES
SSSSVSPKLEUFTFVIVSYALTCLIKARGDPSRPVGYGFAVDCRSLMNPVSPVS
SYTFETFMEEGFALAAARVSDSEALDENVALKIPLELFGFTLSPGTQVLVSAGSTR
FGVYGLDFGWRPERKVVVYSDQGEAISFAESRDGSGGVGLFSLKKHEMDVLVDLLH
KGLN"
complement(join(27466..27879,28046..29073))
/note="gene_id:MXF12.8"
/codon_start=1
/pseudo
/evidence=not_experimental
/product="transposase-like protein"
complement(join(30804..30874,30947..31289))
/note="contains similarity to Ac-like transposase"
gene_id:MXF12.9"
/codon_start=1
/pseudo
```

```
CDS
/evidence=not_experimental
32432..33823
/note="gene_id:MXF12.10"
/codon_start=1
/evidence=not_experimental
/product="anthocyanin acyltransferase-like protein"
/protein_id="BAB10830.1"
/db_xref="GI:10177551"
/translation="MVSSLNIEVARVTPSNSDSSAESLSPLTYFDLIYVKLRAVERV
IFYRITNTRPFEDSVIIVPNLKTLSLCLSHLPLAGKLWPELDLPKPKIVYSONDDV
SFSVAETNADFSSLSGNEPFTSTELIPLVPAQSSDSSASIVSQVTLFPNQGFCIGV
SAHVAVLGKTEPFTTELYPLVPELHVSDSASAVSFQVTLFNPQSENKLVNWR
SFTKILAAGKEPANPKSLKLLWSPGDPVVRTYLTQLTREDIQKRLERLKRVSSESS
STSSKELRLSTFVIVSYVLCIIRANGGEPRHPVAFVSDCRSLMNPPTNPFYFGN
CIAGCSRMLAKTAKMEGEGLLAAATVSDSLEENDESFAWKIPDFVATVLPPELQ
ILVSSNRNGYVELDFGWRPKVMVVISPONGISMAESRDONGSVEIGFSLKKHEM
DPLDILLHCELFU"
34363..35709
Query Match 4.8%; Score 72.6; DB 12; Length 66237;
Best Local Similarity 48.6%; Pred. No. 5.1e-06;
Matches 228; Conservative 0; Mismatches 239; Indels 2; Gaps 1;
QY 899 GTTCATTAATCACCATTGATGTGCGATCTCTCTTGTATCCACCAATTTCCACACGCTACTT 958
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26298 GTACGGATTTCGCTGTGCAITGTCGAAGCCTTATGGTTCCACCGGTTCCATCGAGTTATTT 26357
QY 959 TGGTAACCTGGTGCACCATGTGTC--CCGACCTTAAAAAATGCTGTTTGACTACGGAA 1016
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26358 TGGTAATGTGATCTGCTTGTAAATAATGTCGTTAACGGCAGAGACTTTTATGAGTGAA 26417
QY 1017 AATGGGTATGCACTTGGTGCTAAAGTAATGGAGAGTCTATATGCAAAATGATATATAAT 1076
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26418 GAAGGGTTTTAGTCTCTCGGAGAATGGTTAGTGATCTGTTGAGGCATTTGGATGAGAAT 26477
QY 1077 AAGCGGGAATCTTGAAGATGCCGAGATGCCGATGCATGCAATCTTTCATGATCCCGGCTAGG 1136
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26478 GTTGCAATGAAGATTCAGAGATTTTGGAAGGGTTTACGACTCTTTCACCGAGGAACACAG 26537
QY 1137 AAGATTGGTGTCTGCTGACACCTACCACTTCTAGACATTTTACGATTTTGGTGGGGG 1196
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 26538 GTTTTCTGTTTCCCGGGTGCACCGGTTGCGAGTTTACGGGTGGGTTTGGGTGGGGT 26597
QY 1197 AAGCGCATAAAGTATGAGACTGTTTCAATAGACTATAAATACGTCGATTTCTATAAATGCA 1256
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26598 AGACCCGAGAAGTGTGGTGTGTCAAATTGACCAAGGTGAACGGATTTCTTTTCCGGAG 26657
QY 1257 AGCAAAACATCAGCACAGAATCTTGAAATTTGGAATTTGAGTCTACCGAGTATGCAAAATGGAG 1316
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26658 AGTAGAGATGGGAGTGGTGTGTAGACTTGGCTTCTCCCTCAAGAAACATGAATGGAT 26717
QY 1317 GCCTTTTCTAGCATCTTGTGATGAAGGATTAGAGAGTCAAGATTTTCATTTGT 1365
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26718 GTTCTTTGTTGATTGTTTCAAGAGGACTAGAAAAATTAATCAATTTT 26766

RESULT 12
AX083744 LOCUS AX083744 1141 bp DNA PAT 28-FEB-2001
DEFINITION Sequence 22 from Patent WO0111061.
ACCESSION AX083744
VERSION AX083744.1 GI:13185472
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 1141)
AUTHORS Kunst,L. and Clemens,S.
TITLE Regulation of embryonic transcription in plants
JOURNAL Patent: WO 011061-A 22 15-FEB-2001;
FEATURES
Location/Qualifiers
1..1141
```

```
/organism="synthetic construct"
/db_xref="taxon:32630"
1. 1141
/note="consensus sequence of A.t., L.a., and B.n. FAEL
promoters"
BASE COUNT 123 a 32 c 42 g 112 t 832 others
ORIGIN

Query Match 4.0%; Score 60; DB 10; Length 1141;
Best Local Similarity 12.4%; Pred. No. 0.0034;
Matches 105; Conservative 275; Mismatches 467; Indels 0; Gaps 0;

Qy 52 ATCGCTCGTACTCTTCTTCTCGACATTACTTGTGCTACTCTCTCCCTCGGTCACC 111
Db 292 AKYWGNNABVNSTCTTWSKTKTKVTSWANNCRAGDANKDKHKKWSAAGVYWN 351
Qy 112 ATCTTTTCTTATGACTTTCACATCTTAAATCCATTTTCATGGACACTATTGTTCCCA 171
Db 352 NNNWTKKARHBDWVWHSAAKKWHANAHHYSRKKWTBYKRTMVNNNGTTTMMKRW 411
Qy 172 GCCTAAACAATCTTTATCGGTCACCTCTCAACATTTTTCCTGTTTCCGTAGTAATTTGA 231
Db 412 AWYKXNDMDWBGTYNNNNNGRTYTGWNKKMWTYYKKKANNCKWRADHKTCETHNNT 471
Qy 232 TTGATTTCTTCAACATGATGCTCGGCTTTTAAATAAAAAACACAGAAATGACGTTG 291
Db 472 WKKMKTNNWCKWKSMTNGKSHRBAAYVYTWMMWRYYAHANNNDYWWKACWYKYB 531
Qy 292 AAGGTGATCTGTTGTTGTTACTTTTTCGAGAAATGTTGCTTCACTTAAATTTGACAG 351
Db 532 VCSKWNNAAYAAVTKSSNNYTSRYRWKTNNSWRSDTRSGRANNYARABHYGKWT 591
Qy 352 GAAATCATCTCCGAAATGCGAACTTTTATCCACTTGTACCTTCATTTGGAAATGCA 411
Db 592 RMBWSHTWBHBRAGAHHWBMWYBAKCHCKAWYKAKYAGAGSSNNNNNNNNNN 651
Qy 412 TCAAAATATGATGCTGCTCAGGCTCCCACTTTTTCCTCACTCAAGTACGTTTTCGG 471
Db 652 NNNATCARDYYASRWYAMANKWYKBAANNAYYHANNWGCWNNATDTRTMKN 711
Qy 472 GCTCGGATATACATGAGATGAGCAATCATCATAGCTTGGTGACGCTAGCAGCGGT 531
Db 712 NNNNAGATWKNNNNNNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAKNAK 771
Qy 532 TCAACTTTTGAAGGGTGGACTTCGATTATTCATCTGGTGTAGATCGTCTTTTAA 591
Db 772 RNNGAYTKYTTNNNNYRGVVTNTAARDGANNNNNNNNNNNNNNNNNNNNNNNN 831
Qy 592 CGAAAGGATCTCCACCGGTTTGTATAGATTGATTAACTCCCATTTAGATGAAATA 651
Db 832 YGTNNNNNNNNNAYAWNTNKWYTTDDRWBRBAYTNNNNNNNNNNNNNNNNNNNN 891
Qy 652 AGTTCAGACATACAGGCTCGAAAGTTTATTAACCTTCGACGCTTGTGTCCTCAGT 711
Db 892 CDAMKWDATKMNNAATYNRGTAWRTNNNNNNNNNNNNNNNNNNNNNNNNNNNN 951
Qy 712 ATAAAGTTCGCTCAAGTTTGTGTGACCGCACTTAATCACTACTAAGAAAAGG 771
Db 952 WVKATKTKTGGWNNCTTTCRRYKNNCTWYTWMTTTRTWTAAATRWKNTNNATGSMTRCNA 1011
Qy 772 TCTTAACCCAGTGCACCAACTTGGAGTACATGCTCATCTTTTACGGTAACCTGTGTTATA 831
Db 1012 TCWKNNNYWTGKTKWTAYRMAATRWKAWKVMATGSWNTNSYARWAYKTRAYKGYNA 1071
Qy 832 TATGAGTTCGATACGCAAACTACCTGCTGTAATAAATAGGAGAAGAAAGCGGCAAGCAGT 891
Db 1072 CAWRWRWKGATCYMTDANWWTACATSWMATHYKNHMKCNKNNNNNNNNNNNNNN 1131
Qy 892 TAGACA 898
Db 1132 CDGARYW 1138
```

```
RESULT 13
CNS06V71/c 805 bp DNA STS 11-JAN-2001
LOCUS T7 end of clone AX0AA017D07 of library AX0AA from strain CBS 7064
DEFINITION of Pichia farinosa, sequence tagged site.
ACCESSION AL416788
VERSION AL416788.1 GI:12197692
KEYWORDS STS
SOURCE Pichia farinosa.
ORGANISM Pichia farinosa
Eukaryota; Fungi; Ascomycota; Saccharomycotina: Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Pichia.
REFERENCE 1 (bases 1 to 805)
AUTHORS de Montigny J., Spehner, C., Souciet, J., Tekala, F., Dujon, B.,
Wincker, P., Artiguenave, F. and Potier, S.
TITLE Genomic Exploration of the Hemiascomycetous Yeasts: 15. Pichia
sorbitophila
JOURNAL FEBS Lett. 487 (1), 87-90 (2000)
PUBMED 11152890
REFERENCE 2 (bases 1 to 805)
AUTHORS Souciet, J.L., Aigle, M., Artiguenave, F., Blandin, G.,
Bolotin-Fukuhara, M., Bon, E., Brottier, P., Casaregola, S.,
de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,
Malpertuy, A., Neveuglise, C., Ozier-Kalogeropoulos, O., Potier, S.,
Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M.,
Wincker, P. and Weissenbach, J.
TITLE Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of
yeast species for molecular evolution studies(1)
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
PUBMED 11152876
REFERENCE 3 (bases 1 to 805)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This STS is part of a random genomic sequencing program of thirteen
yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
5 kb were prepared and both extremities were sequenced. See
keywords for description of this sequence and for the sequence of
the other extremity of this insert.
FEATURES
source
1. 805
/organism="Pichia farinosa"
/strain="CBS 7064"
/db_xref="taxon:4920"
/clone="AX0AA017D07"
/clone_lib="AX0AA"
/note="end : T7"
1. 805
STS 310 a 108 c 62 g 238 t 87 others
BASE COUNT
ORIGIN

Query Match 3.7%; Score 55.8; DB 53; Length 805;
Best Local Similarity 38.8%; Pred. No. 0.029;
Matches 108; Conservative 41; Mismatches 129; Indels 0; Gaps 0;

Qy 1231 ATAATACGTCGATTTCTATAAATGCAACAAACATCAGCACAGATCTTGAATTGGAT 1290
Db 471 ATAAAAAGATTTTGCAATCAATGAGGATATTATTGAGGAAAKKGGTTTTTTTTTTTWT 412
Qy 1291 TGAGTCTACCGATGATCAAAATGGAGCGTTTCTTCTAGCATCTTTCATGAAGATTAGAGA 1350
Db 411 WTTGCTTTAKGAWATTTTACTTGWATTTTATKTTWTTTTAKGAGAGTAATTTTGW 352
Qy 1351 GTCAAGTTTCATGTTAGATCATCGTCCCTTTTGTGTGTCATCAAGTTTCTGTCGTTTT 1410
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:11:47 ; Search time 445.49 seconds
(without alignments)
2125.470 Million cell updates/sec

Title: US-08-894-356c-5

Perfect score: 1508

Sequence: 1 TGAACATTCGCAACATGCC.....AAAAAAAAAAAAAAAAAAAA 1508

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

1: /SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1500	99.5	1508	17	AAT37312 Aromatic acyl tran
2	205.6	13.6	1703	17	AAT37308 Aromatic acyl tran
3	185.8	12.3	1479	17	AAT37311 Aromatic acyl tran
4	166.6	11.0	1518	17	AAT37313 Aromatic acyl tran
5	136.6	9.1	1622	17	AAT37309 Aromatic acyl tran
6	121.6	8.1	936	22	AAF58252 Oligonucleotide D1
7	121.6	8.1	936	22	AAF58254 Oligonucleotide D1
8	121.6	8.1	936	22	AAF58257 Oligonucleotide D1
9	121.6	8.1	936	22	AAF58259 Oligonucleotide D1
10	121.6	8.1	936	22	AAF58262 Oligonucleotide D2
11	121.6	8.1	938	22	AAF58255 Oligonucleotide D1

12	118.2	7.8	936	22	AAF58252	Oligonucleotide D1
13	118.2	7.8	936	22	AAF58254	Oligonucleotide D1
14	118.2	7.8	936	22	AAF58257	Oligonucleotide D1
15	118.2	7.8	936	22	AAF58259	Oligonucleotide D2
16	118.2	7.8	936	22	AAF58262	Oligonucleotide D1
17	118.2	7.8	938	22	AAF58255	Oligonucleotide D1
18	86	5.7	1648	21	AAAC64785	Lemon acyl transe
19	53.8	3.6	1379	21	AAC46852	Arabidopsis thalia
20	52.6	3.5	244	22	AAF58238	Oligonucleotide D1
21	50.8	3.4	860	21	AAAC74426	Human secreted pro
22	48.2	3.2	1469	21	AAAC95491	Human secreted pro
23	48.2	3.2	1925	21	AAAC78138	Human cancer assoc
24	48	3.2	244	22	AAF58238	Oligonucleotide D1
25	47	3.1	1285	21	AAAC59828	Human secreted pro
26	47	3.1	1875	19	AAV33017	Dirofilaria immiti
27	47	3.1	1875	19	AAV33017	Dirofilaria immiti
28	46.6	3.1	1534	17	AAT15995	Human interleukin-
29	46.2	3.1	4391	18	AAT60666	Mouse CD100 antige
30	45.8	3.0	3184	21	AAAC75894	Human ORFX ORF1449
31	45.2	3.0	1140	21	AAAC98004	Human colon cancer
32	45.2	3.0	1377	22	AAAF44479	Mouse dextran sodi
33	45.2	3.0	1974	21	AAAS4282	Human mesenchymal
34	45	3.0	1122	21	AAAC59379	Human secreted pro
35	44.8	3.0	259	18	AAV00423	3' fragment of clo
36	44.8	3.0	1694	20	AAAX07431	Homo sapiens secre
37	44.8	3.0	2907	21	AAAC98917	Human pancreatic c
38	44.6	3.0	637	21	AAAC79045	Human secreted pro
39	44.6	3.0	858	21	AAAC59226	Human secreted pro
40	44.6	3.0	1641	17	AAT15758	Human interleukin-
41	44.4	2.9	2163	21	AAAF21924	Human breast and o
42	44.2	2.9	3495	20	AAAC33814	Coding sequence fo
43	44.2	2.9	3521	14	AAQA1226	Clone GP3 encoding
44	44	2.9	1694	19	AAV33190	Secreted protein B
45	44	2.9	2497	21	AAAC61197	Human cDNA encodin

ALIGNMENTS

RESULT 1
AAT37312
ID AAT37312 standard; cDNA to mRNA; 1508 BP.
XX
AC AAT37312;
XX
DT 06-FEB-1997 (first entry)
XX
DE Aromatic acyl transferase coding sequence.
XX
KW Aromatic acyl transferase; transformation; anthocyanin pigment;
KW plants; acylation; colour; tone; colouration; colour change;
KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;
KW Senecio cruentus; Lavandula angustifolia; ds.
XX
OS Senecio cruentus (Clone pCat8).
XX
FH Key Location/Qualifiers
FT CDS 3..1367
FT /tag= a
FT /product= Aromatic acyl transferase.
XX
PN WO9625500-A1.
XX
PD 22-AUG-1996.
XX
PF 16-FEB-1996; 96WO-JP00348.
XX
PR 30-JAN-1996; 96JP-0046534.
PR 17-FEB-1995; 95JP-0067159.
PR 29-JUN-1995; 95JP-0196915.
XX
XX (SUNR) SUNTORY LTD.
XX

PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
 PI Nakao M, Tanaka Y, Yonekura K;
 XX WPI; 1996-393401/39.
 DR P-PSDB; AAW04726.

PT DNA coding for aromatic acyl transferase - for transforming plants
 PT which produce anthocyanin pigments and thus altering colour tone,
 PT e.g. of flowers

PS Claim 4; Page 69-72; 94pp; Japanese.

CC Vectors containing DNA fragments encoding proteins of plant origin
 CC with aromatic acyl transferase activity may be used to transform
 CC plants which produce anthocyanin pigments. The aromatic acyl
 CC transferase acylates the pigments in the flower resulting in colour
 CC tone changes and allowing new colourations to be produced. Six
 CC specific DNA sequences encoding aromatic acyl transferase from
 CC different plants are described in AAT37308-T37313.

XX Sequence 1508 BP; 442 A; 294 C; 295 G; 477 T; 0 other;

Query Match 99.5%; Score 1500; DB 17; Length 1508;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1503; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TGAACATTCTCGAACATCCCGCAATATCGGCCCTCGGCACCATCGGCCATCGCTCGT 60
 DB 1 tgaacattctcgaaatcccgcaaatatcgccctcgggcaaccatcgccatcgctcgt 60
 QY 61 TATCTCTTACTTCTCGACATTACTTGGCTACTTCCCTCCGGTCACCATCTTTTCT 120
 DB 61 tatctcttacttctcgacattacttggctacttctccctccggctccaccatctttct 120
 QY 121 TCTATGACTTTCACATTTCTAAATCCCATTTTCATGGACACTATTGTTCCCGAGCTAAAC 180
 DB 121 tctatgacttctcacatTTCTAAATCCCATTTTCATGGACACTATTGTTCCCGAGCTAAAC 180
 QY 181 AATCTTTATCGGCTACTCTTCAACATTTTTCGCCGTTTTCCTAGTAGTAATTTGATTTTC 240
 DB 181 aatctttatcggtcactcttcaacatTTTTCGCCGTTTTCCTAGTAGTAATTTGATTTTC 240
 QY 241 CTAACACGTGATGGTTCGGTTTTAAATAAAACACAGAAATAAACACGTTGAAGTGATT 300
 DB 241 ctaacacgtgatgggttcgggtTTTAAATAAAACACAGAAATAAACACGTTGAAGTGATT 300
 QY 301 CTGTTGTGTTACTTTTCAGAAATGTTCTGTGCTTTTAAATAATTTGACAGAAATATC 360
 DB 301 ctgttgtggttactTTTTCAGAAATGTTCTGTGCTTTTAAATAATTTGACAGAAATATC 360
 QY 361 CTCGAAATGTGAACATTTTATCCACATTTGACCTTCAATTTGGGAAATGCAATCAATPAT 420
 DB 361 ctcgaaatgtgaaacTTTATCCACATTTGACCTTCAATTTGGGAAATGCAATCAATPAT 420
 QY 421 GTGATTCGCTCAGCTCCACATTTTTCACATTTCAAGTCAAGTCAAGTCAAGTCAAGT 480
 DB 421 gtgatTCGCTCAGCTCCACATTTTTCACATTTCAAGTCAAGTCAAGTCAAGTCAAGT 480
 QY 481 TATCACTAGGAATGACGAATCATATAGCTTGGTGGCTAGCAGCGGTTCAACTTTT 540
 DB 481 tatcactaggaaTGACGAATCATATAGCTTGGTGGCTAGCAGCGGTTCAACTTTT 540
 QY 541 TGAAGGTTGACTTCGATTTATCAATTCGTTGATAGTCGCTCTTTTAAACGAAAGAT 600
 DB 541 tgaaggTTGACTTCGATTTATCAATTCGTTGATAGTCGCTCTTTTAAACGAAAGAT 600
 QY 601 CTCACCGGTTTTCATAGATTGATTACATCCACATTTTATAGATGAATTAAGTTGACAC 660
 DB 601 ctccacggTTTTCATAGATTGATTACATCCACATTTTATAGATGAATTAAGTTGACAC 660
 QY 661 ATACAAGCTCGAAAGTTTATATAACCTTCGAGCGCTTGTGGTCCCACTGATAAAGTTC 720
 DB 661 atacaagCTCGAAAGTTTATATAACCTTCGAGCGCTTGTGGTCCCACTGATAAAGTTC 720

DB 661 atacaaggctcgaaagttttttataaaacctcgagccttgttgggtcccaactgataaagttc 720
 QY 721 GGTCAACGTTTGTGTGACCGCACTAATATCAATCTACTAAAGAAAGTCTCTTAACCC 780
 DB 721 ggtcaacgTTTGTGTGACCGCACTAATATCAATCTACTAAAGAAAGTCTCTTAACCC 780
 QY 781 AAGTGCCTAACTTGGAGTACATCTTTTACGGTAACTTGTGGTTATATATGGAATTT 840
 DB 781 aagtgcctaaactTGGAGTACATCTTTTACGGTAACTTGTGGTTATATATGGAATTT 840
 QY 841 GCATAGCGAAATCACTCGTAAATAATAGGAGAAAGGCGAAGACCGATTTAGAACAGT 900
 DB 841 gcatagcgaaatCACTCGTAAATAATAGGAGAAAGGCGAAGACCGATTTAGAACAGT 900
 QY 901 TCATAATCACCATTGATTGTGATCTCGTCTTGTATCCACCAATTTCCACAGCTACTTTG 960
 DB 901 tcataatcaccattGATTGTGATCTCGTCTTGTATCCACCAATTTCCACAGCTACTTTG 960
 QY 961 GTAACGTGTGTGACCATGTGTCCGACCTTTAAANAATGTCGTTTGTGACTACGGAAATG 1020
 DB 961 gtaactGTGTGACCATGTGTCCGACCTTTAAANAATGTCGTTTGTGACTACGGAAATG 1020
 QY 1021 GGTATGCACTTGGTCTAAAGTAAATTTGGAGAGTCTATATGCAAAATGATATATAAAGG 1080
 DB 1021 ggtatgcaCTTGGTCTAAAGTAAATTTGGAGAGTCTATATGCAAAATGATATATAAAGG 1080
 QY 1081 ACGGAATCTTGAAGAATGCCGAGATGCGATGAACCTTTTCATGATCCCGCTAGGAAGA 1140
 DB 1081 acggaatctTGAAGAATGCCGAGATGCGATGAACCTTTTCATGATCCCGCTAGGAAGA 1140
 QY 1141 TTGGTGTGCTGGTACACCTAAGCTCAACTTGTAGGACTTTTGATTTGGGTGGGGAAGC 1200
 DB 1141 ttggtGTGCTGGTACACCTAAGCTCAACTTGTAGGACTTTTGATTTGGGTGGGGAAGC 1200
 QY 1201 GCATAAGATATGAGACTGTTTCAATAGACTATAATACCTGATTTCTATAAATGCAAGCA 1260
 DB 1201 gcataaGATATGAGactgtttcaatagactataaTAcgtatctctataaataGcaagca 1260
 QY 1261 AAACATCAGCACAAGATCTTGAAATTTGGATTGGAGTCTACCGAGTATGCAATTTGGAGCGT 1320
 DB 1261 aaacatcagcaGaatcttgaaattggattggactcaccgagtatgcaaatggagcggt 1320
 QY 1321 TTTCTAGCATCTTTCATGAAGGATAGAGTCAAGTTTCATTTAGTTCCTTCTATTCTTTCCT 1380
 DB 1321 ttctagcatctTTTCATGAAGGATAGAGTCAAGTTTCATTTAGTTCCTTCTATTCTTTCCT 1380
 QY 1381 TTTTGTGTGCTCAAGTTTCTGCTGTTTTTATGAGTTGCCACTGTTCTTATTCTTTAAGTA 1440
 DB 1381 ttttGTGTGCTCAAGTTTCTGCTGTTTTTATGAGTTGCCACTGTTCTTATTCTTTAAGTA 1440
 QY 1441 TACCTTTGCTACTATGTTTGAAGATGCAACGATATAAATGAAATAAATAAATAAATAA 1500
 DB 1441 tacctttGCTACTATGTTTGAAGATGCAACGATATAAATGAAATAAATAAATAAATAA 1500
 QY 1501 AAAAAAAA 1508
 DB 1501 aaaaaaaa 1508

RESULT 2

AAT37308
 ID AAT37308 standard; cDNA to mRNA; 1703 BP.

XX AAT37308;

XX AC
 DX 06-FEB-1997 (first entry)

XX Aromatic acyl transferase coding sequence.

XX Aromatic acyl transferase; transformation; anthocyanin pigment;
 KW plants; acylation; colour; tone; colouration; colour change;
 KW Gentiana triflora; Petunia hybrida; perilla ocimoides;
 KW Senecio cruentus; Lavandula angustifolia; ds.

XX OS Gentiana triflora var. japonica (Clone pGAT4).

XX FH Key Location/Qualifiers

FT CDS 6..1415

FT /*tag= a

FT /product= Aromatic acyl transferase.

XX PN W09625500-A1.

XX PD 22-AUG-1996.

XX PF 16-FEB-1996; 96WO-JP00348.

XX PR 30-JAN-1996; 96JP-0046534.

XX PR 17-FEB-1995; 95JP-0067159.

XX PR 29-JUN-1995; 95JP-0196915.

XX PA (SUNR) SUNTORY LTD.

XX PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;

XX PI Nakao M, Tanaka Y, Yonekura K;

XX WIPI; 1996-393401/39.

XX DR P-PSDB; AAW04722.

XX CC DNA coding for aromatic acyl transferase - for transforming plants which produce anthocyanin pigments and thus altering colour tone, e.g. of flowers

XX PS Claim 4; Page 53-57; 94pp; Japanese.

XX CC Vectors containing DNA fragments encoding proteins of plant origin with aromatic acyl transferase activity may be used to transform plants which produce anthocyanin pigments. The aromatic acyl transferase acylates the pigments in the flower resulting in colour tone changes and allowing new colourations to be produced. Six specific DNA sequences encoding aromatic acyl transferase from CC different plants are described in AAT37308-T37313.

XX SX Sequence 1703 BP; 512 A; 353 C; 356 G; 482 T; 0 other;

Query Match 13.6%; Score 205.6; DB 17; Length 1703;

Best Local Similarity 52.2%; Pred. No. 3.7e-39;

Matches 722; Conservative 0; Mismatches 614; Indels 48; Gaps 10;

QY 1 TGAACATTCTCGAACATGCCGAATATCGGCCCCCTCGGGCACCA---TCGGCCATCGCT 57

DB 25 tgaaggctcttgaaaaatgccaaagtacaccaccatctgacacaacagatgctcagattat 84

QY 58 CGTTATCTCTTACTTCTTTCGACATTACTTGGCTACTCTTCCCTCCGGTCCACCATCTTT 117

DB 85 cgtaccggttaacattcttcgatataccccgggttgacattgaaataagatgcagtccttc 144

QY 118 TCTTCTATGACTTTCACATTTCTAAATCCCATTTCTATGGACACTATTGTTCCCGAGCTAA 177

DB 145 tcttttaacagcttccgttaaccacaagaacattctctggacactgtataccctaactcta 204

QY 178 ACAATCTTTATGGTCACTCTTTCACATTTTTCCTCCGGTTGCTAGTAATTTGATTTGAT 237

DB 205 agccctcttctctcactctctaaacactacctctccgtcttagcggaatttgtatgc 264

QY 238 TTCCTAACACTGATGTTCTCGGGTTTTTAATAAANAACCAAAATAAACACAGTTGAAGGTG 297

DB 265 cgaatca-----atcgggcgaaatgccgaagtttcagttactcccgctgatgaggcg 315

QY 298 ATTCTCTTGTGGTACTTTTGGCAGAATGTGCTTGACTTTTAATAAATTTGACAGAAATC 357

DB 316 actcgataactttgatcgtttgcygagctctgaccagagattttgactacccttaagggtcatc 375

QY 358 ATCCTCGAANAATGAAAACTTTTATCCACTTTGTACCTTTCATTGGGAAATGCAATCAAT 417

DB 376 aactggtagattccaatgatttgcatggccttttttatgtatgctgcacgggttataaaga 435

QY 418 ---TATGTGATTGCGTCACSGTCCCACTTTTTCACTTCAAGTGACGTTTTCCTCGGGCT 474

DB 436 ccatgcaagactataaaagtgtaccgcgtctgtagcogtgcgaataaccgttttctcctaacc 495

QY 475 CGGTATATCATAGGAATGACGAATCATATAGCCCTTGGTACGCTAGCAGCGGTTCAC 534

DB 496 gtggcatagccgtggtctctgacggcacatcattcaatgcagatgctaaaaagtttggtaa 555

QY 535 ACTTTTGTAAAGGGTGGAGCTTCGATTATTCATTCGATGATGATGATGATGATGATGATG 594

DB 556 tgttcataatgcttgggacctataatttaacaaatttgggaagacgoggaacttgttccg 615

QY 595 AAGGATCTCCACCGGTTTGTATAGATTTCATTAACA-----TCCACATTTACATGA 646

DB 616 cgaatctctccatctcttcgatagatcataatcaaatctgtatggccctagagagaa 675

QY 647 AATAAGTTTGAGACATACAAAG---GCTCGAAAGTTTATTAACCTTCGAGCCTTTGTTG 702

DB 676 cattttggaacgaatgcgaagatgttcttgaaatgttctctagatttggaaagcaaacccc 735

QY 703 GTCCCACTGATAAAGTTCGCTCAACGTTTGTGTGACCCGAACATAATATCAATCTACTAA 762

DB 736 ctcgattcaacaaggtaacgagctacatatgtcctctccttgcgtaaatccagaagctaa 795

QY 763 AGAAAAGCTCTTAA---CCCAAGTGCCANAATTTGGAGTACATGT--CATCTTTTACGG 816

DB 796 agaacaagtactgaatctcagaggatccgaacgaataacgtgtaacgaagttcacaa 855

QY 817 TAACTTGTGTTATATATGGAGTTGCATAGCCGAAATCA-----CTCGTAAAAAATAG 867

DB 856 tgacgtgtgatacgtatggacatgcattgcataacaaagatgcgtatcagagag 915

QY 868 GAGAAAGAAGGGCGAAGAGAGATTAGAACAGTTTCAATATCACCATTGATGTGCGATCTC 927

DB 916 aatcatcgacacgacgaaatgagctcagctactctcttgcagcggatgctcgagagac 975

QY 928 GTCTTCATCCCAATTTCCACAGCCCTACTTTGGTAACCTGGTGGTGCACCATGTGCCCGA 987

DB 976 tctgacgcccccgctgcgcctaaactcttggcaactctctgcgtcatcgctgcgca 1035

QY 988 CCTTAAAAAATGCTGTTTGTGACTACGGAATGGGTATGCACCTTGTGCTTAAAGTAATTG 1047

DB 1036 aagcaacacataaaagagttagtgtgggataaaaggctctctgttcagttgcagctatcg 1095

QY 1048 GAGAGTCTTATGCAAAATATATATAAATAGGACCGGAATCTTTGAAAGATGCCCGAGAT 1107

DB 1096 gagaagccattgaaaagaggttgcaacgaaaaagcgctcttgcagatgcaaaaactt 1155

QY 1108 GG--CATGAACCTTTTCATGATCCCGGCTAGGAAGATT--GGTGTGCTGGTACACCTA 1161

DB 1156 ggtaacggaacttaattggaatcccttcaaaagattcttcgggatacccggtatcgcccta 1215

QY 1162 AGCTCAACTTGTACGACTTTTATTTTGGTGGGGAAGCGCATATAAGATATGAGACTGTTT 1221

DB 1216 agtcgattcgtatggtatgtatttggatgggaagcctgcataaatttgacattacct 1275

QY 1222 CAATAGACTATATAGCTGATTTCTTATTAATGCAAGCAAAACATCAGCACAGATCTTG 1281

DB 1276 cgtgtgattatgcagaattgatttatgtattcagtcagggttttggaaaaaggtgtg 1335

QY 1282 AATTCGATTGAGTCTACCGNGTATGCAATGAGGCGGTTTCTAGCATCTTTGATGAAG 1341

DB 1336 agattggagtaactatgccttaagattcatdggatgcttgcataaaactcttgaagag 1395

QY 1342 GATT 1345

DB 1396 gctt 1399

RESULT 3
AAT37311

QY 1132 CTA---GGAAGATTGGTGTGCTGTACACCTAAGCTCAACTTGACGACTTTGATTTTG 1188
 Db 1123 aaactgttatttctcgtggcgagatcgagcagcttgatctttacgcgcgattttg 1182
 QY 1189 GTGGGGGAGCGCATAAAGTAGACTGTTTCAATAGACTATATACGT---CGATTT 1245
 Db 1183 gatggggtaagcgggtgaagcaagagactctgctgattgagagagatttaccagatg 1242
 QY 1246 CTAATAATGCAAGCAACATCAGCACAGATCTTGAATTCGATTGATCTACCGAGTA 1305
 Db 1243 cgttctgtaaacccgagggatcgctcgagagattggaggttgattctcttgcacaagg 1302
 QY 1306 TCAAAATGAGCGGTTTCTAGCATCTTTGATGAAGGATTAGAGAGTCAAGTTTCATTGT 1365
 Db 1303 aggaattgcaagcttttgatgattatttttgcgagggagataaagggttgatttaacattt 1362
 QY 1366 AGATCATGTCCTCCCTTTTGTGTGATCAAGTTTCTGCTGTTTATGAGTTCGACATGT 1425
 Db 1363 a-----atcatgtattatgaagttggatgaaatcctctgtttcatctctattgtt 1412
 QY 1426 TCTATTCTTTAAGTATACCTTTTCGACTATGTTTTCAGAGATGCAACGATATAAATGAANA 1485
 Db 1413 taacaataatttttccattgaaacttttttgagccaataaaaaaaaaaaaaaaaaa 1472
 QY 1486 AAAAAAAAAAAAA 1498
 Db 1473 aaaaatgaaaaa 1485

RESULT 5

AAT37309
 ID AAT37309 standard; cDNA to mRNA; 1622 BP.
 AC AAT37309;
 XX
 DT 06-FEB-1997 (first entry)
 DE
 XX Aromatic acyl transferase coding sequence.
 KW Aromatic acyl transferase; transformation; anthocyanin pigment;
 KW plants; acylation; colour; tone; colouration; colour change;
 KW Gentiana triflora; Petunia hybrida; Perilla oclimoides;
 KW Scenecio cruentus; Lavandula angustifolia; ds.
 XX
 OS Gentiana triflora var. japonica (Clone pGAT106).
 PH
 FT Key Location/Qualifiers
 CDS 6..1415
 FT /*tag= a
 FT /product= Aromatic acyl transferase.
 FT 701..703
 FT misc_feature
 FT /*tag= b
 FT /transl_except= ATC encodes Leu
 FT 1322..1324
 FT /*tag= c
 FT /transl_except= AAT encodes Asp
 XX
 XX WO9625500-A1.
 XX
 XX 22-AUG-1996.
 XX
 XX 16-FEB-1996; 96WO-JP00348.
 XX
 XX 30-JAN-1996; 96JP-0046534.
 PR 17-FEB-1995; 95JP-0067159.
 PR 29-JUN-1995; 95JP-0196915.
 XX
 XX (SUNR) SUNTORY LTD.
 XX
 XX Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
 PI Nakao M, Tanaka Y, Yonekura K;
 XX

DR WPI; 1996-393401/39.
 XX P-PSDB; AAW04723.
 PT DNA coding for aromatic acyl transferase - for transforming plants
 FT which produce anthocyanin pigments and thus altering colour tone,
 XX e.g. of flowers
 PS Claim 4; Page 57-61; 94pp; Japanese.
 XX
 CC Vectors containing DNA fragments encoding proteins of plant origin
 CC with aromatic acyl transferase activity may be used to transform
 CC plants which produce anthocyanin pigments. The aromatic acyl
 CC transferase acylates the pigments in the flower resulting in colour
 CC tone changes and allowing new colourations to be produced. Six
 CC specific DNA sequences encoding aromatic acyl transferase from
 CC different plants are described in AAT37308-T37313.
 XX
 SQ Sequence 1622 BP; 478 A; 354 C; 313 G; 477 T; 0 other;
 Query Match 9.1%; Score 136.6; DB 17; Length 1622;
 Best Local Similarity 48.5%; Pred. No. 6e-23;
 Matches 675; Conservative 0; Mismatches 644; Indels 72; Gaps 8;
 QY 1 TCAACATTCTCGAATCCCGCAATATCGGCCCTCGGCACCATCGCCATCGCTCGT 60
 Db 57 tcaagttcttgagaaatgcccgtgtgtgcgccaccacgcgcgtccgagtttacag 116
 QY 61 TATCTCTTACTTCTTCGACATTACTTGGCTACTCTTCCCTCCGGTCCACCATCTTTTCT 120
 Db 117 tccactgtcgttttccgacatgcgatggttgatctgtatgcagaacaccatctgcatt 176
 QY 121 TCTATGACTTTCCACATTCTAAATCCCATTTTCATGGACACATATTTCCCGAGGTAAAC 180
 Db 177 tctacagattccgcacattctgtcccaactctaaatttatcttccatccca---ttaa 233
 QY 181 AATCTTTATCGGTCTACTCTTCAACATTTTTTCCCGTTTGTCTAGTAATTTGATTTATTC 240
 Db 234 cgtcccttccctgttctcaaacacttttccgttagccggaatttggattggccgg 293
 QY 241 CTACACGTGATGGTTCGGGTTTTTAATAAAACCAGAAATAAACACACGTTTGAAGTGATT 300
 Db 294 tagattctctccg-----atagaatgcggagtgctgttacaagaaggggact 341
 QY 301 CTGTTGTGGTTACTTTTGCAGAATGTTCTGTGACTTTTAATAATTTGACAGAAATCATC 360
 Db 342 cgtttcttacaattgcagaatccagatggatttggattatctcgcgagagatc 401
 QY 361 CTCGAAATGTGAACATTTTATCCACTTTGACCTTCATTTGGGAAATGCAATCAATTTAT 420
 Db 402 agagggtattctataaaattcaacgatttgcctccagctcccaagaccattgtaacct 461
 QY 421 GTGATTCGCTCAGGTCCTCCACTTTTTCACCTCAAGTCACGCTTTTTCGGGCTCGGTA 480
 Db 462 ccggcagcgaagattaccacttttgccttccaggtgacggtgtcttcccaaccggta 521
 QY 481 TATCATTAGGAATGACGAATCATCATGCTTGGTGGCTAGCAGCGGTTCACACTTTT 540
 Db 522 tatgcattggacgaatctctcatcaagtcttggtagcaggttcttcttgcatttta 581
 QY 541 TGAAGGCTGACTTCGATTATTCAAATCTGGTGTAGAT-----CGGT 582
 Db 582 ataatatgggttttgggtgacaaatccaatggagattcattaaagtctcttccacttt 641
 QY 583 CTTTTTTAACGAAAGGATCTCCACCGGTTTTTGTAGATTGATTAAACATCCACATTTAG 642
 Db 642 ctctctacctatgtacacagatctgtggtgcaagatcccatcttccattctcgtcgaaaaa 701
 QY 643 ATGAAATAGTTGAGACATACAAAGGCTCGAAAGTTTTTATAAACCCTTCGAGCCTTGTG 702
 Db 702 tctacaatgaaagaaactgctcaaatctcagggcacactactgttcttaaatccagaa 761
 QY 703 GTCCACTGATAAAGTTTCGGTCAACGTTTGTGTGACCCGCAACTAATATCAATCTACTAA 762


```

PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX Umek RM;
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
XX Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
XX Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;
SQ

Query Match 8.1%; Score 121.6; DB 22; Length 938;
Best Local Similarity 0.9%; Pred. No. 1.7e-19;
Matches 7; Conservative 474; Mismatches 283; Indels 0; Gaps 0;

QY 114 CTTTTCTCTATGACTTTCACATCTTAATCCCATTTTCATGACACTATTTGTTCCAGG 173
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 788 CWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 729

QY 174 CTAAACAATCTTATCGGTCACTCTTCAACATTTTTCGCCGTTTCTAGTAATTTGATT 233
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 728 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 669

QY 234 GTATTTCCTAACACTGTTCCGGTCTTTTAATAAAACCAGAAATAAACACGTTGAA 293
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 668 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 609

QY 294 GGTGATCTGTTGTTACTTTTGAGAAATGTTGTTGACATTTAATAATTTGACAGGA 353
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 608 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 549

QY 354 AATCATCTCGAAATGTGAAACATTTTATCCACTTTGTACCTTCATGGGAATGCAATC 413
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 548 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 489

QY 414 AAATTATGTGATTGCGTCCACGTCCTCTTTTTCACCTCAAGTGACGTTTTCGCCGGC 473
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 488 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 429

QY 474 TCGGGTATATCACTAGGAATGAGAAATCATCATGAGCTTTGGTGACGCTAGCACGCGTTC 533
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 428 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 369

QY 534 AACTTTTGAAGGGTGGACTTCGATTATCAATCTGTTGATGATCGGTCCTTTTAAAG 593
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 368 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 309

QY 594 AAAGGATCTCCACCGTCTTTTGATAGATTGATTAAACATCCACATTTAGATGAAATAAG 653
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 308 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 249

QY 654 TTGAGACATACAGGCTCGAAAGCTTTTATAAACCTTCGAGCCTTGTGTCCCACTGAT 713
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 248 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 189

QY 714 AAAGTTCGGTCAACGTTTGTGTTGACCCGAACTAATATCAATCTACTAAGAAAGGTC 773
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 188 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 129

```

```

QY 774 TTAACCCCAAGTGCCAAACTTGAGGTACATGTCTTTTACGTAACATCTGTGTTATATA 833
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 69

QY 834 TGGAGTTGCATACGGAATCACTCGTAAATAATAGGAGAAAGAAA 877
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 68 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 25

RESULT 12
AAF58252
ID AAF58252 standard; DNA; 936 BP.
XX
XX AAF58252;
XX
XX 24-APR-2001 (first entry)
XX
XX Oligonucleotide D1835.
XX
XX Electron-transfer group. ETM; mismatch; genotyping;
KW gene expression; ss.
XX
XX Synthetic.
XX
XX WO200107665-A2.
XX
XX 01-FEB-2001.
XX
XX 26-JUL-2000; 2000WO-US20476.
XX
XX 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
XX
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
XX Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
XX Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;
SQ

Query Match 7.8%; Score 118.2; DB 22; Length 936;
Best Local Similarity 0.8%; Pred. No. 1.1e-18;
Matches 6; Conservative 469; Mismatches 282; Indels 0; Gaps 0;

QY 112 ATCTTTCTCTATGACTTTCACATTTTCAATCCCATTTTCATGACACTATTTGTTCCCA 171
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 32 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 91

QY 172 GGCTAAACAATCTTTATCGGTCACTCTTCAACATTTTTCGCCGTTTGTAGTAATTTGA 231
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 92 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 151

QY 232 TTCTATTCTCAACATGATGGTTCCGGTTTTAATAAAACCAGAAATAAACACGTTG 291
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 152 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 211

QY 292 AAGGTGATCTGTGTGGTTTACTTTTGCAGAAATGTTGCTTGTACTTAATAATTTGACAG 351
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```


AC	AAF58257;	
XX		
DT	24-APR-2001 (first entry)	
XX		
DE	Oligonucleotide D1954.	
XX		
KW	Electron-transfer group; ETM; mismatch; genotyping;	
KW	gene expression; ss.	
XX		
OS	Synthetic.	
XX		
PN	WC200107665-A2.	
XX		
PD	01-FEB-2001.	
XX		
PF	26-JUL-2000; 2000WO-US20476.	
XX		
PR	26-JUL-1999; 99US-0145695.	
PR	17-MAR-2000; 2000US-0190259.	
XX		
PA	(CLIN-) CLINICAL MICRO SENSORS INC.	
XX		
PI	Umek RM;	
XX		
DR	WPI; 2001-I59728/16.	

Nucleic acids containing electron-transfer group, useful as labels in PT hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface -

Example 6; Page 127; 159pp; English.

The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETW) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.

Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

Query Match	7.8%	Score 118.2	DB 22	Length 936
Best Local Similarity	0.8%	Pred. No. 1.1e-18		
Matches	6	Conservative 469	Mismatches 282	Indels 0
Gaps				
Qy 112	ATCTTTTCTTCTATGACCTTCCACATTTCTAAATCCCATTTTCATGGACACTATGTGCCCA	171		
Db 32	ATCTTTTCTTCTATGACCTTCCACATTTCTAAATCCCATTTTCATGGACACTATGTGCCCA	171		
Qy 172	GGCTAAACAACATCTTTATCGGTGCTCTCAACATTTTTTCCGGTTTGCTAGTAATTTGA	231		
Db 92	GGCTAAACAACATCTTTATCGGTGCTCTCAACATTTTTTCCGGTTTGCTAGTAATTTGA	231		
Qy 232	TTGTATTCTTAACACTGATGGTTTCGGGTTTAAATAAACAACAGAAATAAACAACGCTTG	291		
Db 152	TTGTATTCTTAACACTGATGGTTTCGGGTTTAAATAAACAACAGAAATAAACAACGCTTG	291		
Qy 292	AAGTGATTCGTGTGTGTTACTTTTGAGAAATGTTGTCTTGACTTTAATAATTTGACAG	351		
Db 212	AAGTGATTCGTGTGTGTTACTTTTGAGAAATGTTGTCTTGACTTTAATAATTTGACAG	351		
Qy 352	GAATCATCTCGGAAATGTGAAACTTTTATCCACTTGTACCTTATCGGGAATGCAC	411		
Db 272	GAATCATCTCGGAAATGTGAAACTTTTATCCACTTGTACCTTATCGGGAATGCAC	411		
Qy 412	TCAAATTATGTGATTCGCTCACGGTCCCACTTTTTCACCTTCAAGTGACGTTTTTTCCGG	471		
Db 332	TCAAATTATGTGATTCGCTCACGGTCCCACTTTTTCACCTTCAAGTGACGTTTTTTCCGG	471		
Qy 472	GCTCGGTATATCACTAGGAATGACGAATCATCATACGCTTGTGACGCTAGCACCGGT	531		
Db 392	GCTCGGTATATCACTAGGAATGACGAATCATCATACGCTTGTGACGCTAGCACCGGT	531		
Qy 452	gctcawww	511		
Db 452	gctcawww	511		

Qy	532	TCAACTTTTGAAGGGTGACCTTCGATTATCAATCTGGTGTAGATCGGCTCTTTTAA	591
Db	452	#####	511
Qy	592	CGAAGGATCTCCACCGTTTTCATAGATTGATTACATCCACACATTTAGATGAAATA	651
Db	512	#####	571
Qy	652	AGTTGAGACATACAAGGCTCGAAGTTTTTATAAACCTTCGAGCCCTGTGTGGTCCCACTG	711
Db	572	#####	631
Qy	712	ATAAGTTCGGTCAACGTTGTGTGTGACCGCAACTAATATCAATCTACTAAGAAAAAGG	771
Db	632	#####	691
Qy	772	TCCTTAACCAAGTCCAAACTTGGAGTACATGTCATCTTTACGGTAACCTTGTGGTTATA	831
Db	692	#####	751
Qy	832	TATGAGTTGCATAGCGAAATCACTCGTAAAAATAGG	868
Db	752	#####	788

RESULT 15
AAF58259
ID AAF58259 standard; DNA: 936 BP.

XX	AAF58259;
AC	
XX	24-APR-2001 (first entry)
DT	
XX	Oligonucleotide D2004.
DE	
XX	Electron-transfer group; ETM; mismatch; genotyping;
KW	gene expression; ss.
XX	Synthetic.
OS	

XX	WO200107665-A2.
PN	
XX	
XX	01-FEB-2001.
PD	
XX	
XX	26-JUL-2000; 2000WO-US20476.
PF	
XX	
PR	26-JUL-1999; 99US-0145695.
XX	
XX	17-MAR-2000; 2000US-0190259.
PA	(CLIN-) CLINICAL MICRO SENS
XX	
XX	Umex RM;
PI	
XX	
XX	WPI: 2001-159728/16.
DR	

XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
PS Example 6: Page 138; 159pp; English.

The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.
 Sequence 936 BP: 6 A; 138 C; 8 G; 8 T; 776 other;

Query Match 7.8%; Score 118.2; DB 22; Length 936;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:04:45 ; Search time 168.74 Seconds
(without alignments)
1691.841 Million cell updates/sec

Title: US-08-894-356C-5
Perfect score: 1508
Sequence: 1 TGAACATTCGCAACATGCC.....AAAAAAAAAAAAAAAAAAAA 1508

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53.6	3.6	7218	1	US-08-232-463-14 Sequence 14, Appl
2	47	3.1	1472	4	US-08-781-420-10 Sequence 10, Appl
3	47	3.1	1472	4	US-08-781-420-12 Sequence 12, Appl
4	46.6	3.1	1534	1	US-08-300-903A-6 Sequence 6, Appl
5	44.6	3.0	1641	1	US-08-300-903A-8 Sequence 8, Appl
6	43.4	2.9	1493	6	5340934-5 Patent No. 5340934
7	42.8	2.8	1098	3	US-09-248-335-35 Sequence 35, Appl
8	42.6	2.8	2589	4	US-08-569-749-1 Sequence 1, Appl
9	42.6	2.8	2589	5	PCT-US96-12860-1 Sequence 1, Appl
10	42.2	2.8	1223	3	US-09-154-874-4 Sequence 4, Appl
11	42	2.8	1474	4	US-08-821-994-64 Sequence 64, Appl
12	41.8	2.8	991	3	US-08-924-747-25 Sequence 25, Appl
13	41.8	2.8	991	3	US-09-247-373B-25 Sequence 25, Appl
14	41.8	2.8	991	4	US-09-296-715-25 Sequence 25, Appl
15	41.8	2.8	1700	2	US-08-897-340-4 Sequence 4, Appl
16	41.8	2.8	1700	3	US-09-252-329-4 Sequence 4, Appl
17	41.6	2.8	893	1	US-08-276-452A-66 Sequence 66, Appl
18	41.6	2.8	893	2	US-08-798-744-66 Sequence 66, Appl
19	41.6	2.8	2030	3	US-08-706-216-3 Sequence 3, Appl
20	41.6	2.8	3319	3	US-08-335-844A-15 Sequence 15, Appl
21	41.4	2.7	2744	3	US-09-071-101-1 Sequence 1, Appl
22	41.4	2.7	2744	3	US-09-369-618-1 Sequence 1, Appl
23	41.4	2.7	2744	3	US-09-369-617-1 Sequence 1, Appl
24	41	2.7	1325	1	US-08-306-691B-51 Sequence 51, Appl
25	41	2.7	1325	2	US-08-464-517-1 Sequence 1, Appl
26	41	2.7	1325	2	US-08-246-361A-1 Sequence 1, Appl
27	41	2.7	1325	3	US-08-463-772-1 Sequence 1, Appl

28	41	2.7	1325	5	PCT-US93-05000-1 Sequence 1, Appl
29	41	2.7	1577	4	US-08-821-994-59 Sequence 59, Appl
c 30	40.8	2.7	350	1	US-08-171-385-14 Sequence 14, Appl
c 31	40.8	2.7	350	3	US-08-361-441B-14 Sequence 14, Appl
32	40.8	2.7	1728	3	US-08-983-950-7 Sequence 7, Appl
33	40.8	2.7	1738	2	US-08-379-482A-2 Sequence 2, Appl
34	40.8	2.7	8920	2	US-08-446-855A-1 Sequence 1, Appl
35	40.8	2.7	8920	4	US-09-150-741-1 Sequence 1, Appl
36	40.6	2.7	1166	5	PCT-US96-12129B-1 Sequence 1, Appl
37	40.6	2.7	1582	3	US-08-545-196B-12 Sequence 10, Appl
38	40.6	2.7	1582	3	US-08-545-196B-12 Sequence 12, Appl
39	40.4	2.7	5173	1	US-08-242-677-1 Sequence 1, Appl
40	40.4	2.7	1542	1	US-08-207-904-1 Sequence 1, Appl
41	40.4	2.7	3706	1	US-08-207-904-16 Sequence 16, Appl
c 42	40.2	2.7	340	1	US-08-171-385-27 Sequence 27, Appl
c 43	40.2	2.7	340	3	US-08-361-441B-27 Sequence 27, Appl
44	40	2.7	1075	4	US-08-400-006B-6 Sequence 6, Appl
45	40	2.7	1332	2	US-09-057-762-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pt29pt-fls
US-08-232-463-14

Query Match 3.6%; Score 53.6; DB 1; Length 7218;
 Best Local Similarity 1.9%; Pred. No. 0.00017;
 Matches 5; Conservative 171; Mismatches 90; Indels 0; Gaps 0;

QY 7 TTCGCAACATGCCGATATCGGCCCTCGGGCCACCATCGCTCGTTATCTC 66
 DB 1177 YY 1236

QY 67 TTACTTCTTCGACATTACTTGGCTACTCTTCCCTCCGTCACCATCTTTTCTCTATG 126
 DB 1237 YY 1296

QY 127 ACTTTCACATCTAAATCCCATTTCAFGACACTATTTGCCAGGCTAAACAATCTT 186
 DB 1297 YY 1356

QY 187 TATCGCTCACTCTCAACATTTTCCCGTTTGTAGTAATTTGATTTCTTCTACACA 246
 DB 1357 YY 1416

QY 247 CTGATGTTCCGGTTTAAATAAAAAA 272
 DB 1417 YYYYYYYYYYYYYYYYYYGTACCAAA 1442

RESULT 2
 US-08-781-420-10
 ; Sequence 10, Application US/08781420
 ; Patent No. 6248872
 ; GENERAL INFORMATION:
 ; APPLICANT: Chandrashekar, Ramaswamy
 ; APPLICANT: Mehta, Kapil
 ; TITLE OF INVENTION: Parasitic Nematode Tranylglutaminase
 ; TITLE OF INVENTION: Proteins, Nucleic Acid Molecules and
 ; TITLE OF INVENTION: Uses Thereof
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Carol Talkington Verser, Ph.D.
 ; ADDRESSEE: Hesk Corporation
 ; STREET: 1825 Sharp Point Drive
 ; CITY: Fort Collins
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80525
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: WordPerfect for Windows, Version 7.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08781,420
 ; FILING DATE: December 3, 1996
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Verser, Carol Talkington
 ; REGISTRATION NUMBER: 37,459
 ; REFERENCE/DOCKET NUMBER: HW-2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 970/493-7272
 ; TELEFAX: 970/484-9505
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1472 nucleotides
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 2..1105
 US-08-781-420-10

Query Match 3.1%; Score 47; DB 4; Length 1472;
 Best Local Similarity 56.0%; Pred. No. 0.0048;
 Matches 89; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 1344 TTAGAGAGTCAAGTTTCATTAGATCATCGTCCCTTTTGTGTGCATCAAGTTTCTGT 1403
 DB 1314 TCAGTTCGGAATGTTTATTCGGTTAGCTTAGGCTTTTGTGTTACCTTATGTTACTGT 1373

QY 1404 CGTTTATGAGTTGCCACTGTTCTATTCTTTAAAGTATACCTTTCGACTATGTTTGAAG 1463
 DB 1374 TGTATTGTATTACTATTTTCCCTTGTGTTTAAATTTAAATAAAATTTTGTGGAA 1433

QY 1464 ATGCAACGATATAAATAAATAAATAAATAAATAAATAAATAA 1502
 DB 1434 AA 1472

RESULT 3
 US-08-781-420-12/c
 ; Sequence 12, Application US/08781420
 ; Patent No. 6248872
 ; GENERAL INFORMATION:
 ; APPLICANT: Chandrashekar, Ramaswamy
 ; APPLICANT: Mehta, Kapil
 ; TITLE OF INVENTION: Parasitic Nematode Tranylglutaminase
 ; TITLE OF INVENTION: Proteins, Nucleic Acid Molecules and
 ; TITLE OF INVENTION: Uses Thereof
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Carol Talkington Verser, Ph.D.
 ; ADDRESSEE: Hesk Corporation
 ; STREET: 1825 Sharp Point Drive
 ; CITY: Fort Collins
 ; STATE: Colorado
 ; COUNTRY: USA
 ; ZIP: 80525
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: Windows 95
 ; SOFTWARE: WordPerfect for Windows, Version 7.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08781,420
 ; FILING DATE: December 3, 1996
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Verser, Carol Talkington
 ; REGISTRATION NUMBER: 37,459
 ; REFERENCE/DOCKET NUMBER: HW-2
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 970/493-7272
 ; TELEFAX: 970/484-9505
 ; INFORMATION FOR SEQ ID NO: 12:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1472 nucleotides
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 US-08-781-420-12

Query Match 3.1%; Score 47; DB 4; Length 1472;
 Best Local Similarity 56.0%; Pred. No. 0.0048;
 Matches 89; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 1344 TTAGAGAGTCAAGTTTCATTAGATCATCGTCCCTTTTGTGTGCATCAAGTTTCTGT 1403
 DB 159 TCAGTTCGGAATGTTTATTCGGTTAGCTTAGGCTTTTGTGTTACCTTATGTTACTGT 100

QY 1404 CGTTTATGAGTTGCCACTGTTCTATTCTTTAAAGTATACCTTTCGACTATGTTTGAAG 1463
 DB 99 TGTATTGTATTACTATTTTCCCTTGTGTTTAAATTTAAATAAAATTTTGTGGAA 40

QY 1464 ATCCAAACGATATAAAATGAAAAAAGAAAAAAGAAAAAAGAAAAA 1502
 Db 39 AA 1

RESULT 4
 US-08-300-903A-6
 ; Sequence 6, Application US/08300903A
 ; Patent No. 5591630
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, Dirk M
 ; APPLICANT: Giri, Judith G
 ; TITLE OF INVENTION: Interleukin-15 Receptors
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Immunex Corporation
 ; STREET: 51 University Street
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98101
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Apple Macintosh
 ; OPERATING SYSTEM: Apple Operating System 7.1
 ; SOFTWARE: Microsoft Word for Apple, Version 5.1a
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/300,903A
 ; FILING DATE: 06-SEPTEMBER-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: USSN 08/236,919
 ; FILING DATE: 06-MAY-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Perkins, Patricia Anne
 ; REGISTRATION NUMBER: 34,695
 ; REFERENCE/DOCKET NUMBER: 2822-A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 206-587-0430
 ; TELEFAX: 206-233-0644
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1534 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..753
 ; US-08-300-903A-6

Query Match 3.1%; Score 46.6; DB 1; Length 1534;
 Best Local Similarity 60.8%; Pred. No. 0.0061;
 Matches 93; Conservative 0; Mismatches 59; Indels 1; Gaps 1;
 QY 1357 TTTCATTGTAGATCATCGTCCCTTTTGTGTCATCAAGTTTCTGTCGTTTATGAGT 1416
 Db 1368 TATATTATATCTCTGTCCTATATATCTATATATATATATATATTTGTTGAT 1427
 QY 1417 TGGCACTGTTCTATCTTTTA-AGTATACCTTTTCGACTATGTTTGAAGATGCAACGAT 1475
 Db 1428 TGTACCTTGTATAAACAAAATAAAACATCTATTTTCAATATTTTAAATGCAAAAAA 1487
 QY 1476 AAAATGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1508
 Db 1488 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1520

RESULT 5
 US-08-300-903A-8
 ; Sequence 8, Application US/08300903A
 ; Patent No. 5591630
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, Dirk M
 ; APPLICANT: Giri, Judith G
 ; TITLE OF INVENTION: Interleukin-15 Receptors
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Immunex Corporation
 ; STREET: 51 University Street
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98101
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Apple Macintosh
 ; OPERATING SYSTEM: Apple Operating System 7.1
 ; SOFTWARE: Microsoft Word for Apple, Version 5.1a
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/300,903A
 ; FILING DATE: 06-SEPTEMBER-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: USSN 08/236,919
 ; FILING DATE: 06-MAY-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Perkins, Patricia Anne
 ; REGISTRATION NUMBER: 34,695
 ; REFERENCE/DOCKET NUMBER: 2822-A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 206-587-0430
 ; TELEFAX: 206-233-0644
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1641 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 3..839
 ; US-08-300-903A-8

Query Match 3.0%; Score 44.6; DB 1; Length 1641;
 Best Local Similarity 56.5%; Pred. No. 0.02;
 Matches 83; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
 QY 1362 TTCTAGATCATCGTCCCTTTTGTGTCATCAAGTTTCTGTCGTTTATGAGTCCCA 1421
 Db 1453 TTATATTATATCTCTGTCCTATATATATATATATATATATATATATATTTGTTGACA 1512
 QY 1422 CTGTTCTATTCTTTTAAGTATACCTTTTCGACTATGTTTGAAGATGCAACGATATAAATG 1481
 Db 1513 TTGTACCTTGTATAAACAAAATAAAACATCTATTTTCAATAAAAAAAAAAAAAAAAAA 1572
 QY 1482 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1508
 Db 1573 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1599
 RESULT 6
 5340934-5
 ; Patent No. 5340934
 ; APPLICANT: TERMINE, JOHN D.; YOUNG, MARIAN F.; FISHER, LARRY W.


```

; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1223 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 55..1223
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 55..990
; US-09-154-874-4

```

Db	1107	TTTTTCGGATCTTTGAATGTTGAACGAGCTTTAACTTTATCATTAATTTAAGCTCTGTTGTT	1166
QY	1466	GCAACGATATATAAATGAAAAAATAAAAAAAAAAAAAAAAAAAAAA	1508

Db 1167 TTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1209

US-08-821-994-64
; Sequence 64, Application US/08821994A
; Patent No. 6228643

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```

: FILE REFERENCE: PED J0100
: CURRENT APPLICATION NUMBER: US/08/821,994A
: CURRENT FILING DATE: 1997-03-22
: EARLIER APPLICATION NUMBER: PCT/GB97/00729
: EARLIER FILING DATE: 1997-03-18
: EARLIER APPLICATION NUMBER: GB 9606062.9
: EARLIER FILING DATE: 1996-03-22
: NUMBER OF SEQ ID NOS: 89
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 64
: LENGTH: 1474
: TYPE: DNA
: ORGANISM: Brassica napus
: US-08-821-994-64

Query Match      2.8%; Score 42; DB 4; Length 1474;
Best Local Similarity 57.7%; Pred. No. 0.088;
Matches 75; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 1379 CTTTGTGTCATCAAGTTCTGCTGTTTTATGAGTTGCCACGTGTTCTATTCTTTAAG 1438
Db 1302 ctttcgctcaagaagaagtgattgggtgttctatgattaaagagaagtcatacaaaat 1361

Qy 1439 TATACCTTTTCGACATGTGTTTTGAAGATGCACGATATAAAATGAAAAAAAAAAAAA 1498
Db 1362 gatattctcttcaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1421

Qy 1499 AAAAAAAAAA 1508
Db 1422 aaaaaaaaaa 1431

RESULT 12

```

RESULT 12

```
US-08-924-747-25
; Sequence 25, Application us/08924747
; Patent No. 6063570
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/924,747
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 991 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: SOYBEAN
; IMMEDIATE SOURCE:
; CLONE: SSM.PK0067.G5
US-08-924-747-25

Query Match 2.8%; Score 41.8; DB 3; Length 991;
Best Local Similarity 59.8%; Pred. No. 0.087;
Matches 70; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 1392 TCAAGTTTCTGCGGTTTATGAGTGGCAGTGTCTATTCTTTAAGTATACCTTTTCGAC 1451
Db 861 TTAATTGACTGTGATTGGTGGGTAATGATTTTAACTAAATAAAAGTGTTC 920

Qy 1452 TATGTTTGAAGATGCAAGATATAAAATGAAAAAATAAAAAAAAAAAAAA 1508
Db 921 AGTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 977

RESULT 13
US-09-247-373B-25
; Sequence 25, Application us/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/247,373B
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 991 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: SOYBEAN
; IMMEDIATE SOURCE:
; CLONE: SSM.PK0067.G5
US-09-247-373B-25

Query Match 2.8%; Score 41.8; DB 3; Length 991;
Best Local Similarity 59.8%; Pred. No. 0.087;
Matches 70; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 1392 TCAAGTTTCTGCGGTTTATGAGTGGCAGTGTCTATTCTTTAAGTATACCTTTTCGAC 1451
Db 861 TTAATTGACTGTGATTGGTGGGTAATGATTTTAACTAAATAAAAGTGTTC 920

Qy 1452 TATGTTTGAAGATGCAAGATATAAAATGAAAAAATAAAAAAAAAAAAAA 1508
Db 921 AGTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 977

RESULT 13
US-09-247-373B-25
; Sequence 25, Application us/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/247,373B
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 991 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: SOYBEAN
; IMMEDIATE SOURCE:
; CLONE: SSM.PK0067.G5
US-09-247-373B-25

Query Match 2.8%; Score 41.8; DB 4; Length 991;
Best Local Similarity 59.8%; Pred. No. 0.087;
Matches 70; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 1392 TCAAGTTTCTGCGGTTTATGAGTGGCAGTGTCTATTCTTTAAGTATACCTTTTCGAC 1451
Db 861 TTAATTGACTGTGATTGGTGGGTAATGATTTTAACTAAATAAAAGTGTTC 920

Qy 1452 TATGTTTGAAGATGCAAGATATAAAATGAAAAAATAAAAAAAAAAAAAA 1508
Db 921 AGTTTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 977

RESULT 14
US-09-296-715-25
; Sequence 25, Application us/09296715
; Patent No. 6171839
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/296,715
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 991 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: SOYBEAN
; IMMEDIATE SOURCE:
; CLONE: SSM.PK0067.G5
US-09-296-715-25

Query Match 2.8%; Score 41.8; DB 4; Length 991;
```

	Query Match	2.8%	Score 41.8;	DB 2;	Length 1700;
	Best Local Similarity	66.3%	Pred. No. 0.1;		
	Matches 55;	Conservative 3;	Mismatches 25;	Indels 0;	Gaps 0;
QY	1426	TCTATTCCTTTAAGTATACCTTTCGACATGATGTTTGAAGATCGAAGATATAAATGAAAA	1485		
Db	1579	TTTATTAAAAAGAAACCAATACTGTAAAWTTTAAAAAGCAAGRATTAAAAA	1638		
QY	1486	AAAAAAAAAAAAAAAAAAAAA	1508		
Db	1639	AAAAAAAAAAAAAAAAAAAAA	1661		

Search completed: November 5, 2001, 18:04:56
Job time: 15096 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:00:56 ; Search time 5816.79 Seconds
(without alignments)
2450.648 Million cell updates/sec

Title: US-08-894-356C-5
Perfect score: 1508
Sequence: 1 TGAACATTCCTCGAACATGCC.....AAAAAAAAAAAAAAAAAAAA 1508

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estrol:*
83: em_estrol2:*
84: em_estrol3:*
85: em_estrol4:*
86: em_estrol5:*
87: em_estrol6:*
88: em_estrol7:*
89: em_estrol8:*
90: em_estrol9:*
91: em_estrol10:*
92: em_estrol11:*
93: em_estrol12:*
94: em_estrol13:*
95: em_estrol14:*
96: em_estrol15:*
97: em_estrol16:*
98: em_estrol17:*
99: em_estrol18:*
100: em_estrol19:*
101: em_estrol20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				JOURNAL COMMENT	
Result No.	Score	Query Match	Length DB ID	Description	
1	130.2	8.6	690 155	BG599447	EST504342
2	124.4	8.2	763 155	BG591758	EST499600
3	116.4	7.7	570 142	BE922784	EST426553
4	110.2	7.3	788 164	BE187619	EST336180
5	107.6	7.1	575 142	BE920468	EST424237
6	107.2	7.1	487 118	AW616206	EST307245
7	107	7.1	606 113	AW221049	EST297518
8	105.8	7.0	508 173	BG097054	EST461573
9	103.4	6.9	574 142	BE922746	EST426515
10	103.2	6.8	583 113	AW221050	EST297519
11	99	6.6	591 167	BE434257	EST405335
12	98	6.5	564 142	BE921494	EST425179
13	92.6	6.1	482 144	BF096979	EST400516
14	91.4	6.1	630 155	BG581249	EST482982
15	89.2	5.9	497 122	AW930934	EST356777
16	87.6	5.8	705 114	AW299055	EST305729
17	87.2	5.8	659 155	BG598154	EST496832
18	85.8	5.7	403 143	BF053202	EST438432
19	85	5.6	542 167	BE436185	EST407263
20	84.6	5.6	518 119	AW650280	EST328734
21	82.6	5.5	687 111	AW102336	sd86d06.Y
22	82	5.4	783 155	BG584708	EST486469
23	81.6	5.4	526 166	BE323055	NF001A07P
24	80.2	5.3	708 138	BE660849	7-C8 Gmax
25	79.4	5.3	676 153	BG455847	NF071A12P
26	79.4	5.3	681 151	BF639071	NF079A08P
27	79.2	5.3	648 154	BG457806	NF034C06P
28	78.8	5.2	408 151	BF643876	NF089D05E
29	78.8	5.2	642 153	BG448305	NF070A05E
30	76.6	5.1	603 118	AW616119	EST296884
31	73.8	4.9	654 151	BF642724	NF070B03I
32	73.8	4.9	656 119	AW692527	NF056D05S
33	73.8	4.9	677 118	AW586972	EST318595
34	72.2	4.8	597 120	AW774948	EST334099
35	71.6	4.7	605 30	AV440958	AV440958
36	71.6	4.7	649 151	BF643475	NF004H05E
37	71.6	4.7	694 117	AW560686	EST315734
38	71.6	4.7	829 155	BG584678	EST486439
39	71.2	4.7	411 142	BE922378	EST426147
40	71.2	4.7	414 142	BE921952	EST425721
41	67.6	4.5	433 120	AW756036	sl12h06.Y
42	63.6	4.2	368 142	BE921953	EST425722
43	62.8	4.2	307 142	BE920010	EST423779
44	62	4.1	360 150	BF520206	EST457675
45	62	4.1	456 142	BE923572	EST427341
ALIGNMENTS					
RESULT 1					
LOCUS	BG599447	690 bp	mRNA	EST	12-APR-2001
DEFINITION	EST504342 cSTS Solanum tuberosum cDNA clone cSTS25122 5' sequence, mRNA sequence.				
ACCESSION	BG599447				
VERSION	BG599447.1 GI:13616583				
KEYWORDS	EST.				
SOURCE	potato.				
ORGANISM	Solanum tuberosum				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.				
AUTHORS	1 (bases 1 to 690) van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chienmingo,A., Bougril,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.				
TITLE	Generations of ESTs from sprouting potato eyes				

LOCUS	BC591758	763 bp	mRNA	EST	12-APR-2001
DEFINITION	EST499600 P. infestans-challenged leaf Solanum tuberosum cDNA clone BPL110A14 5' sequence, mRNA sequence.				
ACCESSION	BC591758				
VERSION	BC591758.1 GI:13609898				
KEYWORDS	EST.				
SOURCE	potato.				
ORGANISM	Solanum tuberosum				
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Serrnatoophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.				
AUTHORS	1 (bases 1 to 763) Zhang, P., Hernandez, M., Tornqvist, C.-E., Wirtz, U., Loukolanov, A., Rangel, P., Haberlach, G.T., Cho, J., Chiemingo, A., Bougri, O., Buell, C.R., Ronning, C.M., Helgeson, J. and Baker, B.				
TITLE	Generation of ESTs from Potato Leaves Challenged with Phytophthora infestans, Incompatible Reaction				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Cathy Ronning The Institute for Genomic Research For clone info: please contact Research Genetics, Libraries Division tel 1-800-711-6195, email cdna@resgen.com Seq primer: M13F-R.				

```

seq primer: 1518F-K.
Location/Qualifiers
1. 763
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="BPL110A14"
/clone_lib="P. infestans-challenged leaf"
/tissue_type="leaf"
/dev_stage="6 week old"
/lab_host="SOLR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Whole plants were challenged with 450,000
sporangia/ml P. infestans US-1(US 940501) in Biotron
(Madison, Wisconsin). Leaf tissue was collected at 1, 2,
5, 12, and 24 hours post-challenge and frozen in liquid
nitrogen immediately upon removal. Kennebec plants showed
no signs of HR. Katahdin plants (susceptible to P.
infestans US-1) were used as controls and showed
infection. NOTE: We cannot exclude the possibility that
this sequence is actually derived from Phytophthora rather
than potato."
212 a 162 c 156 g 233 t
BASE COUNT
ORIGIN

```

Query Match	8.2%	Score 124.4;	DB 155;	Length 763;
Best Local Similarity	53.2%;	Pred. No. 1.5e-19;		
Matches 297;	Conservative 0;	Mismatches 246;	Indels 15;	Gaps 1;

QY	64	CTCTTACTTTTCGACATTTACCTGGCTACTCTTCCCTCCGGTCCACCATCTTTTCTTCT	123
Db	48	CTCTTACTTATTTTGATCATGTTTGGTTAGGTTCCACCGGTATACGGCGGATATATCT	107
QY	124	ATGACTTTTCACATCTTAAATCCCATTTTCATGGACACTATTGTGCCAGGCTAAACAAT	183
Db	108	ACAAGCTCCCATTTCCAAACCGGATTTGTTCAAAACATTTATTTCTCTTAAAAATT	167
QY	184	CTTTATCGGTCACCTTCAACATTTTTTCCGGTTGCTAGTAATTTGATGTATTTCCCTA	243
Db	168	CACCTCCCTCACCTCTCAACACCTACACGCCCTTAGCAGGAAATGTTGCTCGTCCACTAG	227
QY	244	ACACTGATGGTTCGGGTTTTTAATAAAACCCAGNAATAAACACGGTTGAAGGATTTCTG	303
Db	228	ATACAAATGG-----ATATCTCGAGTTACGTTATGTGACAGGAAATTTCTG	272
QY	304	TTGTGGTTTACTTTTTGACAGATGTTGTCTTGACTTTTAAATTTTGACAGGAAATCATCCTC	363
Db	273	TATCTGTTACTTTTTTGACACGATATGGAATTTCAATTTATCTATTTGGTGACCATCCGC	332
QY	364	GAATAGTGTGAAACTTTTATCCACTTGTACCTTCAATTTGGGAAATGCAATCAAAATTAATGTG	423

Db	333	GTAATGCTAAAGGATTTTATCACTTTGTTCTTAAGTTAGGGAAACCTTAAGGATGCACCCG	392
Qy	424	ATTGCGTCACCGTCCTTTTTCACCTTCAAGTGACCGTTTTCCTCGGGCTCGGGTATAT	483
Db	393	GGGTCCAATTAGCTCCGGTCTTAGCCATTCAAGTGACACTTTTTCGGAATCTTGGTGAT	452
Qy	484	CACTAGGAATGACGAATCATCATGACCTTGGTGAGCGCTAGCACCGGTTCAACTTTTGA	543
Db	453	CCATTGGTTTTCACTAACCATCATGTTGTTGGTGATGGAGCTACCATACTAGGGTTCATTA	512
Qy	544	AAGGTGACCTTCGATTAATTCAACTGGTGCTAGAGTCGGTCTTTTTCACGAAAGGATCTC	603
Db	513	AGGCGTGGCGCTCTACTCCATATAATTCGGTGAGATGAACAATTTCTTATCGAATGAGCTAA	572
Qy	604	CACCGGTTTTTGATAGAT	621
Db	573	TTCCATTTTATGATAGGT	590
RESULT	3		
LOCUS	BE922784	570 bp	mRNA
DEFINITION	EST426553 potato leaves and petioles Solanum tuberosum cDNA clone	EST	02-OCT-2000
ACCESSION	CSBT22E2	5' sequence, mRNA sequence.	
VERSION	BE922784		
KEYWORDS	BE922784.1	GI:10448860	
SOURCE	EST		
ORGANISM	potato.		
REFERENCE	Solanum tuberosum		
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.		
TITLE	1 (bases 1 to 570)		
JOURNAL	van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J., Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.		
COMMENT	Generation of ESTs from potato leaves and petioles Unpublished (2000)		
FEATURES	Contact: Cathy Ronning The Institute for Genomic Research For clone request: please contact Research Genetics, Libraries Division tel 1-800-771-6195, email cdna@resgen.com. Location/Qualifiers 1..570 /organism="Solanum tuberosum" /cultivar="Kennebec" /db_xref="taxon:4113" /clone="CSBT22E2" /clone_lib="potato leaves and petioles" /tissue_type="leaflets and petioles" /dev_stage="8 weeks old plants" /lab_host="SOLR" /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Tissue was supplied by Dr. Fry (Cornell University). Leaflets and petioles were isolated from 8 week old greenhouse grown plants. The plants were watered and fertilized freely. The tissue was immediately frozen in liquid nitrogen."		
BASE COUNT	139 a	136 c	115 g 180 t
ORIGIN			
Query Match	7.78;	Score 116.4;	DB 142; Length 570;
Best Local Similarity	52.5%;	Pred. No. 1.2e-17;	
Matches	289; Conservative	0; Mismatches 246;	Indels 15; Gaps 1;
Qy	5	CATTCTGCAACATGCCGAATATCGGCCCTCGGGACCATCGGCCATCGCTCGTTATC	64
Db	32	CTTGATAGAGCAATGTCAAGTTCGCCACCTTCGGCGGGCGGACGGAGGTGACACTCCC	91
Qy	65	TCTTACTTCTTGACATTAATCTAGTGGCTACTTCTCCCTCGGTCACCACTCTTTTCTTA	124

Db 92 TCTTACTTATTTTGATCATGTTGGTTAGGTTTCCACGGTATACGCGGATATATTCTA 151
QY 125 TGACTTTCCACATCTTAATCCCATTTGATGGACATATTTGTTCCAGGCTAAACAATC 184
Db 152 CAAGCTCCCAATTTCCAAACCGGATTTGCTTCAAAACATATTTCCCTCTTAAAAATTC 211
QY 185 TTTATCGGTCACCTCTTCAACATTTTTCGCCGTTTGTAGTAATTTTGATTTGATTTCCCTAA 244
Db 212 ACTCTCCCTCACTCTCAACACATATACGCCCTTAGCTGGAAAGGTTGCTCGTCCACATGA 271
QY 245 CACTGATGGTTCGGGTTTTTAATAAAAAACAGAAATAAACAACGTTGAAGGTGATTTCTGT 304
Db 272 TACAAACGG-----ATATCTGAGTTTACATTTATGTGACAGGAGATTTCTGT 316
QY 305 TGTGGTTACTTTTGACAGATGTTGCTTGACCTTTAATAATTTTGACAGGAATCATCTCG 364
Db 317 ATCTGTTACTTTTTCGAGACTGATATGATTTCAATTTATCTATTTGGTGACCATCCGG 376
QY 365 AAAATGTGAAACCTTTTATCCACTTGTACTTTCATTTGGGAAATGCAATCAAAATTTATGTGA 424
Db 377 TAATGCTAAGGATTTTATCACTTTGTTCTAGCTTAGGGAACCTTAAGGATCACCCGG 436
QY 425 TTGCGTCACGGTCCCACTTTTTCACCTCAAGTGACGTTTTTTCGGGGCTCGGGTATATC 484
Db 437 GGTCCAATTAGCTCCGGTCTTAGCCATTTCAAGTGACACTTTTTTCCGGAATCTTGGTGTATC 496
QY 485 ACTAGGAATGACCAATCATCATAGCCCTTGGTGACGCTAGCACGCGGTTCAACTTTTTTGA 544
Db 497 CATTTGGTTCACTAACCACATCATTTGTTGGTGATGGAGTACCATAGTAGGGTTTCATTAG 556
QY 545 AGGGTGGACT 554
Db 557 GGCGTGGCT 566

RESULT 4

BE187619
LOCUS EST336180 KVO Medicago truncatula cDNA clone pkV0-1613, mRNA 22-JUN-2000
DEFINITION BE187619 788 bp mRNA EST
ACCESSION BE187619
VERSION BE187619.1 GI:8666803
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
1 (bases 1 to 788)
VandenBosch,K., Endre,G., Hur,J., Moore,J., Beremand,P., Ellis,L.,
Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and
Fraser,C.M.
ESTs from uninoculated seedling roots of Medicago truncatula
Unpublished (1999)
Contact: VandenBosch K
Department of Biology
Texas A&M University
College Station, TX 77843-3258, USA
Tel: 409 845 7707
Fax: 409 845 2891
Email: kate@mail.bio.tamu.edu
Texas A&M University name:T260795e
TIGR sequence name:MTGAU50TK
More information is available at:
http://chryslie.tamu.edu/medicago
Seq primer: Skmod (CTA GAA CTA gta gAT CC).
Location/Qualifiers
1. .788
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pkV0-1613"
/clone_lib="KV0"

FEATURES

source

/tissue_type="Seedling roots"
/dev_stage="Immediately prior to inoculation with
Sinorhizobium meliloti (0 hour)"
/lab_host="E.coli strain XL0R"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0R cells."
BASE COUNT 237 a 188 c 100 g 262 t 1 others
ORIGIN

Query Match 7 3%; Score 110.2; DB 164; Length 788;
Best Local Similarity 54.4%; Pred. No. 3.9e-16;
Matches 279; Conservative 0; Mismatches 213; Indels 21; Gaps 2;
QY 57 TCCTTATCTCTTACTTTCTCGACATTACTTGGCTACTCTTCCCTCGGTCACCATCTT 116
Db 5 TCCTCCCCCTAAACATCTTTGACATTCCTTGGCTTCTTCTCTCAACCCAAACCTA 64
QY 117 TTCTTCTATGACATTTCCACATTTCTAAATCCCATTTTCATGGACACTATTTGTTCCAGGCTA 176
Db 65 TTTTCTATGATTTTCTCTCATTTCCCATTTTCCCATTTTCCACACCACTTGTCCCAAACTT 124
QY 177 AAACAATCTTTATCGGTGACCTTTCAACATTTTTCCTGTTGCTAGTAATTTGATTGTA 236
Db 125 AAACAATCTCTTTCTCTCACACTTCAACATTTACTTCCCATTTTCCGGTACATTT----- 177
QY 237 TTTCTTAACACTGATGTTTCGGGTTTTTAATAAAAAACCAAGAAATAAAACGTTGAAGGT 296
Db 178 -----CGTTCCTCGCTTGATCTAACCGAACCTCACTCGAATTTACTCTTAAC 226
QY 297 GATTCTGTTGTTACTTTTTCGAGAATGTTGCTTGTGACTTTTAATAATTTGACAGGAAT 356
Db 227 AACTCGGTTTCATTCACGGTCCGAGAGTCAAAATAGGCACTTTTGAACACTTGTGCTCCGAT 286
QY 357 CATCTTCGAAATGTGAAACATTTTATCCACTTTGTACCTTCATTTGGGAATGCAATCAA 416
Db 287 TACTCAAGAGATGTTAATGAATTTCCACCTTTTGGTACCAAAAGTTGCAACA---AATCTTT 343
QY 417 TTATGTGATTGCGTCACGGTCCCACTTTTTCACCTCAAGTGACGTTTTCCTCGGGCTCG 476
Db 344 TCATTTGAAGTAAGAAGTTCCCTTTGTTAGCTATTCAATAACATCATTTCCCAATATAT 403
QY 477 GGTATATCACTAGGAATGACGAATCATATAGCCTTTGGTGACGCTAGCACGCGGTTTCAAC 536
Db 404 GGCTTTTCCATTGGCCTAGCTTTTTCATCATGTTGTTGCTGATGGAAGAACCTTCCACAAT 463
QY 537 TTTTGTGAAGGGTGGACTTCGATTATTTCAATCT 569
Db 464 TTCATCAAAACATGGTCTTTCATATTTGTTCTTCT 496

RESULT 5

BE920468
LOCUS EST424237 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION cSTB6C8 5' sequence, mRNA sequence.
ACCESSION BE920468
VERSION BE920468.1 GI:10416544
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 575)
van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buehl,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.

```

TITLE      Generation of ESTs from potato leaves and petioles
JOURNAL    Unpublished (2000)
COMMENT    Contact: Cathy Ronning
           The Institute for Genomic Research
           For clone request: Please contact Research Genetics, Libraries
           Division tel 1-800-771-6195, email cdna@resgen.com.

FEATURES   Location/Qualifiers
            1..575
            /organism="Solanum tuberosum"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone="CSTB6C8"
            /clone_lib="potato leaves and petioles"
            /tissue_type="leaflets and petioles"
            /dev_stage="8 weeks old plants"
            /lab_host="SOLR"
            /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
            XhoI; Tissue was supplied by Dr. Fry (Cornell University).
            Leaflets and petioles were isolated from 8 week old
            greenhouse grown plants. The plants were watered and
            fertilized freely. The tissue was immediately frozen in
            liquid nitrogen."
            167 a      77 c      139 g      192 t

BASE COUNT      167 a      77 c      139 g      192 t

```

	Query Match	7.1k	Score	107.6k	DB	142k	Length	575k
	Best Local Similarity	53.8k	Pred. No.	1.6e-15				
	Matches	299	Conservative	0	Mismatches	239	Indels	18
	Gaps							
Qy	805	CATCTTTTACGGTAACCTTGGTTATATATGGACTTCATATGACGAATCACTCTGTAATAA	864					
Db	17	CATCTTTTACTGTAAAGAGTCTTATGTATGGACTTGTCTGATATAAAATCAAGGACGGGA	76					
Qy	865	TAGGAGAAAGAAAGGGGGAACGACGAGTTAGAACAG-----TTCATATATCAACCATTTGATT	918					
Db	77	TTGGAGAAGAAATGATCATAGATGATAGTAAATGGAGTCTTTTGGATGTCTGGAGATT	136					
Qy	919	GTGATCTCGTCTTGATGCCACCAATCCACAGCCTACTTTGGTAACTGTGGTGCACCAT	978					
Db	137	TTAGAGCAGCATTCAAATCCACCACCTTCTCAATCTTATTTTGGGAATTTGTAGTTGGGT	196					
Qy	979	GTGTCCCGACCTTAAAAA---ATGCTGTTTGTACTACGGAAAAATGGGTATGCACCTTGGTG	1035					
Db	197	GTGTCAAGATCAATAGGCATGTTGACTTTGGAAAGGAAGGTTTTCGAATTCGGG	256					
Qy	1036	CTAAAGTAATTGGAGAGTCTATATGCAAAATGATATAATAATGAGGACGGAATCTTGAAAG	1095					
Db	257	TGGAATTAATTGGAGAAGTCATTCAAAAAAAATGAAGGATGAAGAATGGGTCTCTAATG	316					
Qy	1096	ATGCCGGAGATGCGATGAACCTTTTCATGATCCCGGTACGAAAGATTTGGTTCCTGGTA	1155					
Db	317	GTGATTCGTTAAAGAGTTGGACAACTGACCTTAATTCGATTTTTTTCGAATTCCTGGAT	376					
Qy	1156	CACCTAAGCTCAACTTGTACGACTTTGATTTGGGTGGGGAAGCGCATAAAGTATGAGA	1215					
Db	377	CGCAAAACATGACTTCTATGCTGCTGATTTTGGATGGGAAGGCCGTGAAGTTAGAAT	436					
Qy	1216	CTGTTTCAATAGACTATAAT-----ACGTTCGATTTCTATAAATGCAAGCAAAACAT	1266					
Db	437	TCATTTCTCTGTGACAAATGATGATGGTGGAAATTTTCGATGCTCTTTAGTAAATCTAAGAAT	496					
Qy	1267	CAGCACAAATCTTGAANAATTGGATTGAGTCTACCGAGTATGCAAAATGAGGCGTTTCTA	1326					
Db	497	TTGATGGAGATTAGAGATTGGTTTGTCTATTACTCAATGAATGCTTTTGTCTG	556					
Qy	1327	GCATCTTTTGTATGAAG	1342					
Db	557	CTATATTCACTCACGG	572					

RESULT 6
AW616206

LOCUS	AW616206	487 bp	mrna	EST	24-MAR-2000							
DEFINITION	EST307245 L. hirsutum trichome, Cornell University Lycopersicon hirsutum cDNA clone cLHTD15 5', mRNA sequence.											
ACCESSION	AW616206											
VERSION	AW616206.1	GI:7322240										
KEYWORDS	EST.											
SOURCE	Lycopersicon hirsutum.											
ORGANISM	Lycopersicon hirsutum											
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.											
AUTHORS	van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J. and Tanksley,S.D.											
TITLE	Generation of ESTs from wild tomato (Lycopersicon hirsutum) trichomes											
JOURNAL	Unpublished (2000)											
COMMENT	Contact: David Frisch Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrisch@CLEMSON.EDU 5 prime sequence.											
FEATURES	Location/Qualifiers											
source	1..487											
	/organism="Lycopersicon hirsutum"											
	/db_xref="taxon:62890"											
	/clone="cLHTD15"											
	/clone_lib="L. hirsutum trichome, Cornell University"											
	/tissue_type="trichome"											
	/dev_stage="mixed stages"											
	/notes="Leaves of various stages were shaken in liquid nitrogen, shearing off trichomes. This procedure yielded a mixture of cells which is highly enriched for trichome likely with minor contaminations of other types of leaf cells"											
BASE COUNT	116 a	117 c	91 g	163 t								
ORIGIN												
Query Match	7.1%; Score 107.2; DB 118; Length 487;											
Best Local Similarity	54.4%; Pred. No. 1.9e-15;											
Matches	248; Conservative	0; Mismatches	193; Indels	15; Gaps	1;							
Qy	64	CTCTTACTTTCTTCGACATTA	CTTGGCTACTCTTCCCTCCGGTCCGACATCTTTTCTTCT	123								
Db	36	CTCTTACTTATTGACCATGTTTGGTTAGGTTTCACCTTATTAGCGGATATATCT	95									
Qy	124	ATGACTTTCACATTCAAATCCCATTTTCATGGACACTATTTGTTCCGAGGTAACAAT	183									
Db	96	ACAAAGCTCTCCATTTCCAAATCCGATTTCGTTCAAAACATTTATTCCTCTTAAAAAT	155									
Qy	184	CTTTATCGGTACACTCTCAACATTTTTTCCGGTTTGCTAGTAATTTGATTGTTATTCCTA	243									
Db	156	CACTTTCCCTCACACTCAACACACTATATGCCCTTAGCCGGAAACGTTGCTGTGCCACTAG	215									
Qy	244	ACACTGATGGTTCCGGTTTAAATAAAACACAGAAATAAAACAGCTTGAAGGTGATCTG	303									
Db	216	ATACAAACGGAT-----ATCCTGAGTTACGTTATGTGACAGGATCTG	260									
Qy	304	TGTGGTTACTTTTTCGAGAAATGTTGCTTGACTTTTAATAATTTGACAGGAATCATCCTC	363									
Db	261	TGTCGTGTAATTTTCTCGAGACTGATATGGATTTCAAATTAATCTCATTTGGCGACCATCCG	320									
Qy	364	GAAATCGAACAACCTTTATCCACTTGTACCTTCATTTGGGAAATGCAATCAAAATATGTG	423									
Db	321	GTAATGCTAAGGATTTTATACATTTTGCTTCAAGTTAGGGGAACCTTAAGGATGCACCCG	380									
Qy	424	ATTCCGTCAGGGTCCCACTTTTTTTCACCTTCAAGTGAGCTTTTTTTTCCGGGTCGGGTATAT	483									

```

Db 381 GGTCCAATTAGCCCGCTTCTAGCCATCAAGTGACACTTTTTCGGAATCTTGGGTAT 440
QY 484 CACTAGGAATGACGAATCATCATGCTTGGTGACG 519
Db 441 CCATTGGTTTCACTAACCATCATGTTGTTGGTGATG 476

RESULT 7
AW221049 606 bp mRNA EST 07-DEC-1999
LOCUS EST297518 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION cDNA clone cLEF3F1, mRNA sequence.
ACCESSION AW221049
VERSION AW221049.1 GI:6532733
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 606)
AUTHORS Alcalá, J., Vrebalov, J., White, R., Matern, A.L., Holt, I.E., Liang, F.,
Upton, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Ronning, C.M.,
Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
TITLE Generation of ESTs from tomato fruit tissue
JOURNAL Unpublished (1999)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu
5 prime sequence.
FEATURES
source
location/Qualifiers
1..606
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF3F1"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
BASE COUNT 147 a 143 c 123 g 193 t
ORIGIN
Query Match 7.1%; Score 107; DB 113; Length 606;
Best Local Similarity 50.4%; Pred. No. 2.2e-15;
Matches 300; Conservative 0; Mismatches 280; Indels 15; Gaps 1;

QY 3 AACATTCTCGAATATCGGCCCTCGGGCACCATCGGCCATCGCTCGTTA 62
Db 10 ACCGTATGAGCAATGTCANGTTGGCCACTCCCGGGCCGACGAGGTGATCTC 69
QY 63 TCTCTTACTTTCTTCGACATTAAGTGGCTACTCTCCCTCCGGTCACCACCTTTTCTTC 122
Db 70 CCTCTTACTTATTGACCATGTTTGGTAGGGTTTCGGCGGTATAGGCGGATATTATT 129
QY 123 TATGACTTTCCACATTTCTAATCCCATTTTCATGGACACTATTGTTCCAGGCTAAACAA 182
Db 130 TACAAGCTCTCCATTTTTCAAACCCGATTTTCGTTCAAAACATATTCTCTCTTAAAAAT 189
QY 183 TCTTTATCGGTCACCTCTTCAACATTTTTTCCCGCTTTGCTAGTAGTAATTTGATTGTTTCT 242

```

```

Db 190 TCACCTCTCCCTCACCTCTCAACACACTATACGCCCTTAGCCGGAACCGTTGCTTGTCCACTA 249
QY 243 ACACCTGATGGTTTCGGGTTTAAATAAAAAACCAAGAAATAAAACACAGTTGAAGTGATTC 302
Db 250 GATCAAAACGGAT-----ATCCTGAGTTCAGTTTATGTGACAGAGATTC 294
QY 303 GTTGTGTTTACTTTTGCAGAAATGTTCTTGTGACTTTTAAATAATTTGACAGGAAATCATCCT 362
Db 295 GTGCTGTTTACTTTTTCAGAGACTGATGAATTTCAATATATCTCATTTGGTGACCATCCG 354
QY 363 CGAAATATGCAAAACCTTTTATCCACTTTGACCTTCAATGCGGAATGCAATCAATTTATGT 422
Db 355 CGTAAGGCTAAGGATTTTATCATCACTTTTGTCTTAAGTTAGGGAACCTAAGGATGCACCG 414
QY 423 GATTGCTCAGCGTCCCACTTTTTCACCTTCAAGTCAAGTCAAGTTCCTCCGGCTGGGTATA 482
Db 415 GGGGTCCAACTAGCCCGCTCTTAGCCATTCAGGTGACACATTTTTCGGAATCTTGGTGTA 474
QY 483 TCACTAGGAATGACGAATCATCATAGCCCTTGGTGACCTAGCACGCGTTTCAACTTTTGG 542
Db 475 TCATTGTTTCACTAACCATCATGTTGTTGGTGAGGAGCTACTATAGCAGGGTTCATT 534
QY 543 AAAGGTGAGCACTCGAATATTCATCTGTTGATAGTCGGTCTTTTAAACGAAAG 597
Db 535 AAGGCGTGGGCTCTACTCCCAAAATTCGGTGGACATGAACAATCTTATCGAATG 589

RESULT 8
BG097054 508 bp mRNA EST 29-JAN-2001
LOCUS EST461573 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION GSTB45P9 5' sequence, mRNA sequence.
ACCESSION BG097054
VERSION BG097054.1 GI:12587089
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 508)
AUTHORS Utterback, T., Hansen, C.L., Doan, B., Bougri, O., Buell, C.R., Ronning,
van der Hoeven, R.S., Bezzerides, J., Holt, I.E., Liang, F., Cho, J.,
C.M., Fry, W.E., Tanksley, S.D. and Baker, B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
source
location/Qualifiers
1..508
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB45P9"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 127 a 130 c 87 g 164 t
ORIGIN
Query Match 7.0%; Score 105.8; DB 173; Length 508;
Best Local Similarity 53.7%; Pred. No. 4.2e-15;

```

```
Matches 252; Conservative 0; Mismatches 202; Indels 15; Gaps 1;
Qy 58 COTTATCTTACTTTCTTCGACATTAATGCTTACTCTTCCCTCCGGTCCACCATCTTT 117
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 51 CGCTCCCTCTAATCTTATTTGATCATATATGTTCCGTTTGGCTATATGCGTCGGAATT 110
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 118 TCTTCTATGACTTTCCACATCTTAATCCCATTTTCATGGACACTATGTTCCAGGCTAA 177
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 111 TATTTCTAAGCTACCAATTTCCAAACTCGAATTTGTTTCAAAACCATTTATCTCTACTTA 170
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 178 AACAATCTTTATCGGCTACTCTTCAACATTTTTCCTCGTTGCTAGTAAATTTGATGAT 237
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 171 AACATCTACTCTCCCTACTCTCAAACTACTACACACCTTACCTGGCAACATGCTTGTG 230
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 238 TTCCCTAACACTGATGTTCCGGGTTTTTAATAAAAAACCAAGAAATAAAACACGTTGAAGGTG 297
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 231 CACTAAATTCGAGTGGTT-----ATCCTGAGTTGCTTATGCTGAGG 275
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 298 ATTCTGTTGTTGTTACTTTTTCGAGAAATGTTGCTTGTGACTTTTAATAATTTGACAGAAATC 357
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 276 ATTCTATATCTGTTACTTTTATTGAGACTGATGAATTTCAATCATCTCATCTGTAACC 335
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 358 ATCCCTCGAAATGTCGAAACTTTTATCCACTTGTACCTTCAATGCGAAATGCAATCAAT 417
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 336 ATCCCTCGAAATGCTAAGGATTTTATCCCTTCAATTCCTCAATTTGGCAACCTTAAGGATG 395
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 418 TATGTGATTTGGCTACGGTCCGACCTTTTTCACCTTCAAGTGACGTTTTCGGGCTCGG 477
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 396 CACCGGGGTCAAACTAGTCCCGGCTTAGCCATTCAGTGACACTTTTTCGGAATCTTG 455
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 478 GTATATCACTAGGAATGACGAATCATCATATAGCCTTTGGTGACGCTAGCAC 526
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 456 GCATATCGGTTGGTTTGTAGTAACCATCATGTCGCTTGTGTGGAATATAC 504
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
BE922746
LOCUS BE922746 574 bp mRNA EST 02-OCT-2000
DEFINITION ESF426515 potato leaves and petioles Solanum tuberosum cDNA clone
cSPB21L12 5' sequence, mRNA sequence.
ACCESSION BE922746
VERSION BE922746.1 GI:10448822
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
REFERENCE 1
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
Generation of ESTs from potato leaves and petioles
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
Location/Qualifiers
1..574
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSPB21L12"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site1: EcoRI; Site2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. the tissue was immediately frozen in
```

```
liquid nitrogen."
BASE COUNT 162 a 123 c 106 g 183 t
ORIGIN
Query Match 6.9%; Score 103.4; DB 142; Length 574;
Best Local Similarity 52.5%; Pred. No. 1.6e-14;
Matches 286; Conservative 0; Mismatches 241; Indels 18; Gaps 2;
Qy 115 TTTTCTCTATGACTTTCCACATCTTAATCCATTTTCATGGACACTATGTTCCCGAGC 174
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 23 TTTTATCTACAGCTACCAATTTCCAACTCGAATTTCTGTTCAAAACCATTTATCTCTACTC 82
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 175 TAAACAAATCTTTATCGGCTACTCTTCAACATTTTTCCTCGTTTCTAGTAAATTTGATTG 234
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 83 TTAACATTTCACTCTCCCTCACTCTCAAACTACACACCTTAGCTGCAACATGCTT 142
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 235 TATTTCTTAACACTGATGTTCCGGGTTTTTAATAAAAAACCAAGAAATAAAACACGTTGAAG 294
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 143 GTCCACTAAATTCGAGTGGTT-----ATCCTGAGTTGGGTTATGTGACTG 187
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 295 GTGATTCCTGTTGTTGCTTACTTTTGCAGAAATGTTGCTTTGACTTTTAATAATTTGACAGAA 354
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 188 GAGATTCCTATATCTGTTACTTTTATTGAGACTGATGAATTTCAATCATCTCATCTGTTGTA 247
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 355 ATCATCTCGAAATGTCGAAACTTTTATCCACTTGTACCTTCAATTTGGGAAATGCAATCA 414
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 248 ACCATCTCGAAATGCTAAGGATTTTATTCCTCAATTT--CCTCAATTTGGCACAACTAAGG 304
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 415 AATATGTTGATGCTGCTACGGTCCGACCTTTTTCCTCAAGTGACGTTTTCCTCGGGCT 474
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 305 ATGCACCGGGGTCAAACCTAGTCCCGGCTTAGGCCATTTCAAGTGACACTTTTTCGGAATC 364
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 475 CGGTTATATCACTAGGAATGACGAATCATCATAGCTTTGGTGACGCTAGCAGCGGTTTCA 534
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 365 TTGGCATATCGTTGGTTTGTAGTAACCATCATGTCGCTTGTGATGGAATACCATCGTGA 424
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 535 ACTTTTGAAGGGTGAGCTTCGATTTATCAATTTGCTGTTAGTAGTCGTTTTCCTTTTAAAGA 594
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 425 AATTCATAAAGAACATGGGGTTTACTCAACAAATTCGGCGGTGATGAACAGTGTCTTAGAGA 484
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 595 AAGNATCTCCACCGGTTTTTGTATAGATTAACATCCACATTTTAGATGAAATAGT 654
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 485 ATGAGTTCAATTCATTTATATAGTCCGTGATAGGCTGATAAAAGACCCTTATGAACAAGGACGA 544
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 655 TGAGA 659
Db      |||
Qy 545 TTATA 549
Db      |||

RESULT 10
AW221050
LOCUS AW221050 583 bp mRNA EST 07-DEC-1999
DEFINITION EST297519 tomato fruit mature green, TAMU Lycopersicon esculentum
cDNA clone cLEF3f3, mRNA sequence.
ACCESSION AW221050
VERSION AW221050.1 GI:6532734
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
REFERENCE 1
AUTHORS Alcatal,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ann,S., Ronning,C.M.,
Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
```

100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.

FEATURES

Location/Qualifiers
1..583
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF323"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLEF - Fruit were tagged at the 1cm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"
142 a 140 c 117 g 184 t

BASE COUNT 142 a 140 c 117 g 184 t
ORIGIN

Query Match 6.8%; Score 103.2; DB 113; Length 583;
Best Local Similarity 50.2%; Pred. No. 1.8e-14;
Matches 295; Conservative 0; Mismatches 278; Indels 15; Gaps 1;

QY 3 ACATCTCGAACATGCCGAATATCGGCCCTCGGGCACCATCGGCATCGTCGTGA 62
DB 10 ACCGTGATTGACATATGTCAGTTGCGGCACCTCGCGGGGCAACGAGGTGATATC 69
QY 63 TCTCTTACTTTCTCGACATTACTGGGTACTCTCCCTCGCGTCCACCATCTTTCTTC 122
DB 70 CTTCTTACTTATTTGACCATGTTGGTTAGGGTTTCGCGGTATGAGCGGATATTTAT 129
QY 123 TATGACTTTCCACATCTTAATCCCATTCATGGACACTATTTGTTCCCAAGGTAAACAA 182
DB 130 TACAAGCTCTCCATTTTCAACCCGATTTGTTTCAAAACATTTATCTCTTAAATAT 189
QY 183 TCTTTATCGGTGACATCTTCAACATTTTTTCCCGTTGCTAGTAATTTGATTTCTT 242
DB 190 TCACCTCTCCCTCACCTCTCAACACTATACGCCCTTAGCGGAAACGTTGCTTGTCCA 249
QY 243 AACACTGATGTTTCGGGTTTTTAATAAAACACAGAAATAAACACGTTCAAGGTGAT 302
DB 250 GATCAAAACGGAT-----ATCCTGAGTTAGCTTATGTGACAGGAGATTTCT 294
QY 303 GTTGTGGTTTACTTTGCAGAAATGTTGCTGTGACTTTAATAATTTGACAGGAAATCAT 362
DB 295 GTGCTGTGTTTTCGAGACTGATATGAATTTCAATTTATCTCATTTGTTGACCATCG 354
QY 363 CGAAATGTGAAACTTTTATCCACTGTGACCTTTCATTTGGGAAATGCAATCAATTTAT 422
DB 355 CTTAAGGCTTAAGGATTTTATCACTTTGTTCCTAAGTTAGGGGAACCTAAGGATGCA 414
QY 423 GATTGGGTACCGTCCCACTTTTTCACATTTCAAGTGACGTTTTTTCCGGGCTCGGGT 482
DB 415 GGGGTCCAACTAGCCCGCTCTTAGCCATTTAGGTGACACTTTTTCGGAATCTTGGT 474
QY 483 TCATAGGAATGACGAATCATCATGCTTGTGAGCTTAGCAGCGGTTCACATTTTGTG 542
DB 475 TCCATTTGTTTTCATCAACCATCATGTTGTTGGTGATGGAATCTACTATACATGTT 534
QY 543 AAGGGTGGAATTCGATTTATCAATCTGGGTAGATCGGCTCTTTTATA 590
DB 535 AAGGGTGGGCTCTACTCCACAATAATACGTGGACATGAACCAATCTTA 582

RESULT 11
BE434257
LOCUS 591 bp mRNA EST 24-JUL-2000

DEFINITION EST405335 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA clone CLEGL5H6, mRNA sequence.
ACCESSION BE434257
VERSION BE434257.1 GI:9432100
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum

REFERENCE 1 (bases 1 to 591)
AUTHORS Alcalá J., Vrebalov J., White R., van der Hoeven R.S., Holt I.E., Liang P., Hansen T.S., Craven M.B., Bowman C.L., Ronning C.M., Nierman W., Fraser C.M., Martin G.B., Giovannoni J.J. and Tanksley S.D.
Generation of ESTs from tomato fruit tissue, breaker stage
Unpublished (2000)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.

TITLE Generation of ESTs from tomato fruit tissue, breaker stage
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.

FEATURES
source
1..591
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEGL5H6"
/clone_lib="tomato breaker fruit, TIGR"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescript SKmCuadapt; Site_1: EcoRI; Site_2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."
159 a 129 c 121 g 182 t

Query Match 6.6%; Score 99; DB 167; Length 591;
Best Local Similarity 51.7%; Pred. No. 1.8e-13;
Matches 262; Conservative 0; Mismatches 230; Indels 15; Gaps 1;

QY 115 TTTTCTCTATGACTTTCACACTTCTAAATCCCATTTTCATGGACACTATTGTTCCCGAGC 174
DB 17 TATATTTTACAGCTCTCCNTTTTCAACCCGATTCGTTCAAAACATATATTCCTCCCTC 76
QY 175 TAAACAATCTTTATCGGTCACTCTTTCAACATTTTTTCCCGTTTGTAGTAATTTGATTG 234
DB 77 TTAATAATTCACCTCTCCCTCACCTCTCAACACTATAGCCCTTAGCGGAAACGTTGCTT 136
QY 235 TATTTCTCAACTGATGTTGCGGTTTTTAATAAAACACAGAAATAAACACGTTGAAG 294
DB 137 GTCCACTAGATACAAACGGAT-----ATCCTGAGTTACGTTATGTGACAG 181
QY 295 GTGATCTCTGTGGTTACTTTTGCAGAAATGTTCTTGCATTTTAATAATTTTCACAGGAA 354
DB 182 GAGATCTCTGTCTGTACTTTTTCGAGACTGATGATTAATTTCAATTTATCTCATTTGGTG 241
QY 355 ATCATCTCGAAAATGTGAAAATTTTATCCACTTTGACCTTTCAATTTGGGAAATGCAATCA 414
DB 242 ACCATCCGCTAAGGCTAAGGATTTTATCACCTTCTTCTTAAGTTAGGGAACCTTAAGG 301
QY 415 AATTATGATGCTGTCACGGTCCCACTTTTTCACCTTCAAGTGACGTTTTCCTCGGGCT 474
DB 302 ATGCACCGGGGTCCCAACTAGTCCCGCTCTTAGCCATTCAGGTGACACTTTTTCGGAATC 361

```

QY 475 CGGTATATCACTAGGAATGACGAATCATCATAGCCTTGGTACGCTAGCAGCGGTTCA 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 TTGGTATCACTATGTTTCACTAACCATCATGTTGTTGGTATGAGCTACTATAGCAG 421
QY 535 ACTTTTGAAGGTTGACCTTCGATTAATCAATCGTGTAGATCGGTCCTTTTAAACGA 594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 GGTTCATTAAGCGGTGGCTCTACTCCACAATTCGGTGGACATGAACAATCTTATCGA 481
QY 595 AAGGATCTCCACCGGTTTTCATAGAT 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 ATGAGCTAATTCACATTTATGATAGGT 508

RESULT 12
BE921494
LOCUS BE921494 564 bp mRNA EST 02-OCT-2000
DEFINITION EST425179 potato leaves and petioles Solanum tuberosum cDNA clone
cSTB13B18 5' sequence, mRNA sequence.
ACCESSION BE921494
VERSION BE921494.1 GI:10447486
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 564)
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Uterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cda@resgen.com.
FEATURES
source 1..564
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB13B18"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 150 a 139 c 97 g 178 t
ORIGIN

Query Match 6.5%; Score 98; DB 142; Length 564;
Best Local Similarity 53.2%; Pred. No. 3.2e-13;
Matches 265; Conservative 0; Mismatches 215; Indels 18; Gaps 2;

QY 58 CGTATCTCTTACTTTCTTGACGATTAATGCGTACTTTCCTCCCGTCCGATCATCTTT 117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74 CGCTCCCTCAACTATTTTATCATATATGTTTGGTATGATGCGTGGATTT 133
QY 118 TCTTCTATGACTTTCCACATCTCAATCCCATTTTCATGGACATATTTGTTCCGAGTAA 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 134 TATTTCTACAGCTACCAATTTCCAAATCGATTTTCGTTCAACCATATTTCTACTCTTA 193
QY 178 AACAACTTTATCGCTCACTCTCAACATTTTTCCTCCGTTTGGCTAGTAATTTGATGTAT 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 194 AACATTCACCTCCCTCACTCTCAACACACTACACACCTTACGTCGGCAACATTCCTGTC 253
QY 238 TTCCTAACACTGATGTTGCGGGTTTTTAATAAAAAACAGAAATAAACACGTTGAAGGTG 297

```

```

Db 254 CACTAAATTCGAGTGGTT-----ATCCTGAGTTCGTTATGCTAGCTGAG 298
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 298 ATTCCTGTTGGTTACTTTTCCACAATGTTGCTCTGACTTTTAATAATTTGACAGGAATC 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 299 ATTCATATCTGTTACTTTTATTGAGACTGATATGAATTTCAATCATCTCATTTGGTAACC 358
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 358 ATCCTCGAAATGTGAACACTTTTATCCACTTGTACCTTCAATTCGGAAATGCAATCAAT 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 ATCCTCGAATGCTAAGGATTTTATCCCAT---CCTCAATGGCACACCTAAGGATG 415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 418 TATGTGATTCGCTCACGTCCTCCACTTTTTCACCTTCAAGTGACGTTTTCCTCGGCTCG 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 416 CACCGGGGTCAAACTAGTCCCGGCTTAGCCATTCAGTCACACTTTTTCGGAATCTTG 475
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 478 GTATATCACTAGGAATGACGANTCATCATAGCCTTGGTGACGCTAGCAGCGGTTCAACT 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 476 GCATATCCGTTGGTTTATGTAACCATCATGTCGCTTGTGTAATACCATGCTGAAT 535
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 538 TTTTGAAGGGTGGACTT 555
    ||| ||| ||| |||
Db 536 TCATAAGAACATGCGGTT 553
    ||| ||| ||| |||

RESULT 13
BF096979
LOCUS BF096979 482 bp mRNA EST 19-OCT-2000
DEFINITION EST400516 tomato nutrient deficient roots Lycopersicon esculentum
cDNA clone cLEW18O20 5' sequence, mRNA sequence.
ACCESSION BF096979
VERSION BF096979.1 GI:10902689
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 482)
AUTHORS van der Hoeven,R.S., Garvin,D.F., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T.S., Ronning,C.M., Craven,M.B., Bowman,C.L.,
Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Giovannoni,J.J.
and Tanksley,S.D.
TITLE Generation of ESTs from tomato nutrient-deficient roots
JOURNAL Unpublished (1999)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu
FEATURES
source Location/Qualifiers
1..482
/organism="Lycopersicon esculentum"
/cultivar="TA492"
/db_xref="taxon:4081"
/clone="cLEW18O20"
/clone_lib="tomato nutrient deficient roots"
/tissue_type="roots"
/dev_stage="5-6 weeks old"
/lab_host="SOLR"
/notes="Vector: pBluescriptSKmCuAdapt; Site_1: 5' EcoRI;
Site_2: 3' XhoI; Roots were harvested from plants grown
under the following deficiencies/stresses: 10 mM Al, Zn,
P, K, Fe.N. mRNA was isolated from individual treatments.
Proportional aliquots of mRNA of each treatment were mixed
and used for library construction."
BASE COUNT 113 a 119 c 93 g 157 t
ORIGIN

Query Match 6.1%; Score 92.6; DB 144; Length 482;

```

```

Best Local Similarity 52.8%; Pred. No. 6.2e-12;
Matches 234; Conservative 0; Mismatches 194; Indels 15; Gaps 1;

QY 64 CTCCTACTTCTTCGACATCTTGGCTACTCTCCCTCCGGTCCACCATCTTCTTCTT 123
Db 50 CTCCTACTTCTTTCGACATCTTGGCTACTCTCCCTCCGGTCCACCATCTTCTTCTT 109
QY 124 ATGACTTTCCACATCTTAATCCCATTTTCATGACACTATTGTTCCAGCGCTAAACAAT 183
Db 110 ACAAGCTCTCCATTTTCAACCCGATTTTCGTTCAAAACATTTATCTCTCTTAAATTT 169
QY 184 CTTTATCGGTCTACTCTTCAACATTTTTCCTCCGTTGCTAGTAAATTTGATTTTCCCTA 243
Db 170 CACTCTCCCTCACTCTCAAAACACTATACGCCCTTAGCGGAAACGTTGCTTGTCCACTAG 229
QY 244 ACACATGATGTTGGGTTTTTAATAAAAAACAGAAATAAACAACGTTGAAGGTGATTCTG 303
Db 230 ATACAAACCGAT-----ATCCTGAGTTAGCTTATGTGACGAGGATCTGTG 274
QY 304 TTGTGTTACTTTTTCGAGATTTGTTCTTGTGACTTTAATAATTTTGACAGGAAATCATCTTC 363
Db 275 TGTCTGTTACTTTTTCGAGACTGATATCAATTTCAATTTATCTATGTTGTGACCATCCGC 334
QY 364 GAAATGTGAAACATTTTATCCACTTTTACCTTTCATTTGGGAAATGCAATCAATTAATGTG 423
Db 335 GTAAGGCTAAGGATTTTATCACTTTGTTCTTAAGTTAGGGGAACCTAAGGATGACACCGG 394
QY 424 ATTGCGTACAGGCTCCACATTTTTCACCTTCAAGTGACGTTTTTTCGGGCTCGGGTATAT 483
Db 395 GGGTCCAACTACGCCGCTTTAGGCAATTCAGGTGACACTTTTTCGGAATCTTGGTGTAT 454
QY 484 CACTAGGAATGACGAATCATCAT 506
Db 455 CCATGTTTCTACTAACCAATCAT 477

```

```

RESULT 14
LOCUS BG581249 630 bp mRNA EST 11-APR-2001
DEFINITION ESF482982 GVN Medicago truncatula cdna clone pgvn-64023 5' end,
mRNA sequence.
ACCESSION BG581249
VERSION BG581249.1 GI:13596313
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
REFERENCE 1 (bases 1 to 630)
AUTHORS Fedorova,M., Pleron,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Town
,C.D., Van Aken,S., Utterback,T., Cho,J. and Fraser,C.M.
ESTs from one month old nitrogen-fixing root nodules of Medicago
truncatula, 2001
JOURNAL Unpublished (2001)
COMMENT Contact: Carroll P. Vance
Department of Agronomy and Plant Genetics
University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St.Paul, MN 55108 USA
Tel: 612 625 5715
Fax: 651-649-5058
Email: vance004@maroon.tc.umn.edu
University of Minnesota name: M382308e TIGR sequence name:
MTCDG98TK More information is available at: http://www.medicago.org
Seq primer: SKmod (CTA gaa CTA gta gaa CC).
Location/Qualifiers
1. .630
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pgvn-64023"
/tissue_type="GVN"
/tissue_type="N2-fixing root nodules"

```

```

/dev_stages="effective root nodules harvested one month
post inoculation with Sinorhizobium meliloti"
/lab_host="E. coli strain XL0LR"
/notes="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
effective root nodules harvested one month post
inoculation with Sinorhizobium meliloti. The cDNA was
directionally ligated into the Uni-ZAP XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-ZAP phage using Ex-Assist
helper phage and propagated in XL0LR cells."
BASE COUNT 186 a 140 c 72 g 232 t
ORIGIN
Query Match 6.1%; Score 91.4; DB 155; Length 630;
Best Local Similarity 53.1%; Pred. No. 1.2e-11;
Matches 265; Conservative 0; Mismatches 201; Indels 33; Gaps 2;
QY 57 TCGTTATCTTACTTCTTCGACATTTACTTGGCTACTCTTCCCTCCGGTCCACCATCTT 116
Db 75 TCATCTCCCTTACTTCTTTCGACATCTATGGTTAAGATTACCACTGTTTCAAGAAT 134
QY 117 TTCTTCTATGACTTTCACATTTCTAAATCCCATTTTCATGACACTATTGTTCCAGGCTA 176
Db 135 TTCTTCTATGACTTTCACACCAACCTCATTTTCTTCAAAACCTCTTCTCCCAACTC 194
QY 177 AAACAATTTTATCGGTCTACTCTTCAACATTTTTCCTCCGTTTCTAGTAAATTTGATTGTA 236
Db 195 AAAAAATCACTTTCCATTCACATTTCTCACTTTTATCTCTTTTGGTCTATCTCAATTTGG 254
QY 237 TTTCTTAACACTGATGTTCTGGTCTTAAATAAAACCAAGAAATAAACAACGTTGAAGGT 296
Db 255 CCTCATGA-----TTCACAAAGCAATCATTAATTCATCAAGGA 296
QY 297 GATTCCTGTTGTTACTTTTCGAGAAATGTTGCTTTCGACTTTTAAATTTTGACAGGAAAT 356
Db 297 AACACTTTTCACTTACCGTAGCTGAATCTCATGCTGATTTCAATCATTTCTCTGGTAAG 356
QY 357 CATCTCGAAAATGTGAAACATTTTATCCACTTGTACCTTTCATTTGGGAAATGCAATCAAA 416
Db 357 AATCTTTCTGAAGCTACACAGATTCATGATTTTACCTAACCTTGAATATTTCATCA----- 410
QY 417 TTATGTGATGTTGCTGACGTTCCACTTTTTCACCTTCAAGTGACGTTTTCCTCCGGCTCG 476
Db 411 -----CATGACCAAGCTCTTATTTGGCTCTTCAAGTTACTATTTCCTCAAACTAT 461
QY 477 GGTATATCACTAGGAATGACGAATCATCATAGCTTGGTGACGCTAGCAGCGGTTTCAAC 536
Db 462 GGATTTTCAATTTGTTATTTACCTCATCATCATGCTGTTTATAGATGTTAAACTTCAACTTCT 521
QY 537 TTTTGTGAAAGGTTGGACTT 555
Db 522 TTTTCAAAATCTTGGGCTT 540

```

```

RESULT 15
LOCUS AW930934 497 bp mRNA EST 30-MAY-2000
DEFINITION EST356777 tomato fruit mature green, TAMU Lycopersicon esculentum
cDNA clone cLEF42M23 5', mRNA sequence.
ACCESSION AW930934
VERSION AW930934.1 GI:8106335
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 497)
AUTHORS Alcalá,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 16:54:14 ; Search time 8904.87 Seconds
(without alignments)
2643.712 Million cell updates/sec

Title: US-08-894-356C-6
Perfect score: 1522
Sequence: 1 NTCACCACTCTCTCGAATC.....TTTTTTTTTTTTTTTTTTTT 1522

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vil:*
59: gb_vil2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rol:*
95: gb_rol2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	% Match	Query Length	DB ID	Description	
1	765.4	50.3	1479	10	E12756	E12756 Perilla oci
2	763.6	50.2	1476	12	AB029340	AB029340 Perilla f
3	193	12.7	1679	12	AB010708	AB010708 Gentiana
4	193	12.7	1703	10	E12753	E12753 Gentianatri
5	172.6	11.3	1508	10	E12757	E12757 Senecio cru
6	153.8	10.1	1622	10	E12754	E12754 Gentianatri
7	153.8	10.1	1622	12	AB026494	AB026494 Gentiana
8	106.4	7.0	66237	12	AB016892	AB016892 Arabidops

[illegible]

Db 543 AAAGGTGACGTTTCGATTAATCAATCTGCTAGATCGCTCTTTT-----AACGAA 596
QY 592 GAATGTTTCCCGTTTTCGACGATCGCTCGTAATATATATCCGCTAAATTTGACACATAT 651
Db 597 GGATCTCCACCGTTTTCGATAGATTGAT---ACATCCACATTTAGATGAATAAG 653
QY 652 TTATGGAACAACCGCGAGAACGTCGGTTGGAAATCGCAGCATCCATCTTTACCGGAGAT 711
Db 654 TTGAGACATACAAGGCTCGAAAGTTTATAAACCTTCGAGCCTTGTGGTCCCACTGAT 713
QY 712 CGGATTCGAGCTACCTACCTTTTCACCCCAATCTGAAATTAAGAAATTTGAAGTTGATT 771
Db 714 AAGTTTCGCTCAACGTTTGTGTGACCCGAACATATCAATCTACTAAGAAAAAGGTC 773
QY 772 CAGAGAAAAGCCCCAATCTAGTTAATCTCTCTCTCGTCGCGATCGCAGCTTATATC 831
Db 774 TTAACCAAGTCCCAACTTTGGAGTACATGTCATCTTTTACGGTAATCTGTGGTTATATA 933
QY 832 TGGACCGGATCGCCAAATCGGTTCGGAGATTACAAAGA---CGTGATGACGACAAACGC 888
Db 834 TGGAGTTCATAGCGAAATCACTCGTAAANATAGAGAAAGAGGCGGAAGCAGATT 893
QY 889 GCTTTCTTTTAAATCCGATCGAATTAAGCCCGCTTTGGATCCGCGCTCCGGGGAAC 948
Db 894 GAACATTCATAATCAACATGATGTCGATCTCGTCTTGATCCACCAATTTCCACAGCC 953
QY 949 TACTTCGGAACGCTATCTATCGTTTCGGATGGCGAAGATCCTCGCGGGATTTGGTCGGA 1008
Db 954 TACTTTGGTAACGTTGGTGCACCATGTGTCGCGACCTTAAAAAATGCTGTTTGTAGTAGC 1013
QY 1009 GATGAAGGGGTGTTTCGGCGAGTCAGGCGATCGCGCGGGAATAGAGAGGAGCAGC 1068
Db 1014 GAAATGGCTATCCACTTGGTCTAAGTAATTTGGAGATCTATATGCAAAATGATATAT 1073
QY 1069 GACAAGAAGATCTAGAAACTGTGGAGAACTGGCCGCTGTAGATTCGCGAAGCCTTGCAA 1128
Db 1074 AATAAGACGGAATCTTGAAGATGCGCGAGATGCGATCAACCTTTTCATGATCCCGGCT 1133
QY 1129 AACTGTTATTTCTCGGTGCGGGATCGAGCAGCTTTGATCTTTACGCGCGGATTTTGA 1188
Db 1134 A---GGAAGATTGGTGTGCTGTACACCTAAGCTCAACTTGTACGACTTGTATTTGGG 1190
QY 1189 TGGGTAAAGCGGTGAAGAAGACTGTCGATTGATGGAGAAAGTTTACGATGCTG 1248
Db 1191 TGGGGGAAGCGGATAAGTATGAGACTGTTTCAATAGACTAATAAGCT---CGATTTCT 1247
QY 1249 TTGTGTAAACCGAGGATGCTCGCGGAGGATTTGGAGGTTGGATGCTTTGGCCAAAGGAG 1308
Db 1248 ATAAATGCAAGCAAAACATCAGCACAAGATCTTGAAATTTGGATTGAGTCTACCGAGTATG 1307
QY 1309 GAATTCGAAGCTTTTCATGATTATTTTCGGAGGGAATAAGGGTTGATTAAATCATTTA- 1367
Db 1308 CAATGGAGCGGTTTCTAGCATCTTTGATGAAGGATTAGAGAGTCAAGTTTCATTTGTAG 1367
QY 1368 -----ATCATGTATTATGAAGTTGGATGAAATCCCTGCTGTTTCATCTCTATTGTTTA 1418
Db 1368 ATCATGCTCCCTTTTGTGTGATCAAGTTCTGTGCTGTTTATGAGTTGCCACTGTTTC 1427
QY 1419 ACAATAATTTTTTCCATGAACTTTTTTGGTCAATTAATAAAAAAATAAAAAA 1478
Db 1428 TATTCTTTTAAGTATACCTTTTCGACTATGTTTGAAGATCAACGATATAAAATGAAAAA 1487
QY 1479 AAATGAAAAA 1489
Db 1488 AAAAAA 1498

RESULT 6
E12754
LOCUS E12754 1622 bp DNA PAT 24-JUN-1998
DEFINITION Gentianatrilflora mRNA for acyltransferase, complete cds.
ACCESSION E12754
VERSION E12754.1 GI:3251586

KEYWORDS JP 1997070290-A/2.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1622)
AUTHORS Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y.,
Yonekura,K., Mizutani,M. and Kusumi,T..
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 2 18-MAR-1997;
SUNTOY LTD
COMMENT OS Gentianatrilflora
PN JP 1997070290-A/2
PD 18-MAR-1997
PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO,
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
C12R1:865),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FT source 1..1622
FT /variety='japonica' /organism='Gentiana triflora' FT
FT /clone='pGAT106'
FT CDS 35..1474
FT /product='acyltransferase'.
FEATURES
source
1..1622
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 478 a 345 c 322 g 477 t
ORIGIN

Query Match 10.1%; Score 153.8; DB 10; Length 1622;
Best Local Similarity 49.4%; Pred. No. 1.2e-21;
Matches 692; Conservative 0; Mismatches 652; Indels 57; Gaps 9;
QY 4 ACCACCTCCTCGAATCCTCCGAGTGGCGCGCTCCAGGACGGTGGCTGAGCAGTCA 63
Db 56 ATCAAAAGTTCTTTGAGAAATGCGCGTGTGCGCCACACCGGACGCCGTCCCGAGTTTACA 115
QY 64 CTCGCCCTCACCTTCTTCGACATGACGTGGCTGCATTTCCACCCCATGCTTTCAGCTTCTC 123
Db 116 GTCCCACTGCTGTTTTTCGACATGCGATGGTTGATCTCTGATCGCAACACCATCTGCAT 175
QY 124 TTCTAGACTCCCTGTTTCZAAACCCGCTTCTCTCGAAACCGCTGTTCCGAAACTCAA 183
Db 176 TTCTACAGATTCGCGCATGCTTGTCCCAACTCTA---AATTTATCATTTTCATCCATTAA 232
QY 184 CAATCTTATCTTAACCTCAACACTTCTTCCCTTTCATGCAATCTTAATCTACCCT 243
Db 233 TCGTCCCTTTCCCTTCTGTTCTCAACACTTTCTTCCCTTACCGGGAAATTTGATTGGCCG 292
QY 244 CTA---TCGCGCGGAGAAATGCGGAGTTCCGGTATCAGAACCGTGACTCGGTTCTTTC 300
Db 293 GTAGATTCCTCCGATAGAAATGCGGAGTTCCGTTTACAAGAAAGGGGACTCCGTTCTTTA 352
QY 301 ACGATTATGGAGTCTCTCGGAGATCATCGGATTCGGCTCATAAATACTACTCTTTGGC 360
Db 353 ACAATTGCGAATCGAGCATGGATTTTGTATTATCTCGCGGAGATCATCAGAGGGATTCT 412
QY 361 CCTAGCGAGATTATGAAGATCTCCAGCTCGCCCGCATATCGAGGAATCTGATCGGAAA 420
Db 413 TATAAATTAACAGATTGTGATTCGCGAGTCCGAGAACCGGATTGTAACCTCCCGCGACGAA 472
QY 421 TTGTTTCAAGTTTTAGCCGCTGCAAGTGACTCTGTTTCCCGGTCGCGGGGTGTCATCGGA 480

Db	473	CTATTACCACTTTTTCGTTTACAGGTGACGGTGTCTCCACACCGGTATATGCAATTGGA	532
Qy	481	ATAACGACGACACACACGTTAGCGATGCTCCATCGTTTGTAGGGTTATGAAGAGTTGG	540
Db	533	CGCAATCTTCATCAAGTCTTGGTGATGCCAGTCTTTTTCGCAATTTTAAATAAATATGG	592
Qy	541	GCTTCCACTCAATAATTCGGAGGAGATGATGAATCTTTGGACGAAAAAGGTGAATGTTTG	600
Db	593	GTTTTGGTTGCAAAATCCAATGGAGATTC---ATTAAAGTTCTTCCACTTCTTCTCTA	649
Qy	601	CCGGTTTTCGACCGATCGCTCGTGAATATTCGCCCTAAATTTGGACACATATTTATFGGAAC	660
Db	650	CTATGTACGACAGATCTGTGTGCAAGATCCATTTTATATTCGTGCGAAAAATCTACAAAT	709
Qy	661	-----AACGGCGCAAAACGTCGTTGGAATC-----GCAGCATCCATCTTTACCGAGC	708
Db	710	GAAGAAAACTGCTCAATCTCAGGGCACACTACTGTCTTAAATCCAGCAATTTCTAAA	769
Qy	709	GATCGGATTCGAGCTACCTACTTTTTCACCAATCTGAATTAAGAAATTAAGAGGTTTG	768
Db	770	GATGAAGTTTCGAGCACCTTCATCTACACCTATTGATATCATGAAGCTCAAGAAATTC	829
Qy	769	ATTCAGAAAAAGCCCCAAATGTA-----GTTAATCTCTCTCTCTCTCTCTCTCTC	813
Db	830	ATTTCTGTCAAAAAATCGCAACTTAACCCGCTAGTAGTAATATTAATCTGTCAACTTTCACG	889
Qy	814	CGATPCGCGAGCTTATCTTGACCGGCATCGCCAACT-----CGTCCGAGATTAC	864
Db	890	GTGACATCTGCACTGATCTGACATGCTGTGCGAATCATTTAGACACCGCTCGTAGAGAG	949
Qy	865	AAAGCGTGGATGACGACAAACCGCTTTCTTTTAAATTCGGATCGATTTAAGGCCGCGT	924
Db	950	AAAGTGAAGAGGATAACATGACGACAACTTATGTGCTTTCATCAACTGCCGACAACT	1009
Qy	925	TGATTCGCGCGCTCCGGGGAACCTACTTCGGAACCTGCTATCGTTTTCGATGCGCGAAG	984
Db	1010	TTTGCTCCGCGGATACCTTCAAAATTTACTTTTGAATTTGCATAGTGCCTTGTATGGTGGGA	1069
Qy	985	ATCCTCGCGCGGATTTGCTCGGAGATGAAGGGTGTTCGCGGAGCTGAGCGGATCGCG	1044
Db	1070	TCGACTCATGAGCAACTTGTAGGAAATGAGGGTGTTCGCTAGCTGCAACCGCCATCGGA	1129
Qy	1045	GCGGAAATAGAGAGAGGACGAGCGACAAGAA---GATTTAGAACTGTGAGAACTGG	1101
Db	1130	GATGCTATCCATAAGAGGTTTACATGACTACGAAAGAAATCTGAGAGGAGATTTGGATATCG	1189
Qy	1102	CCGTCTGAGATTCGGAAGCCTTGCAAAACCT-----GTTATTTCTCGGTGCGGGA	1152
Db	1190	CGGCCCGGATCAACATCTCGCGCACCAAGGTCGACGCTCATTTATGCTGTTGGATCCGCA	1249
Qy	1153	TCGAGCGGCTTGATCTTTACGGCGCGGATTTTGGATGGGTAAGCGGCTGAAGCAAGAG	1212
Db	1250	CAACGCATATGTCATGATTTTGTATGTCAGATTTTGGTGGGAAAGCTTGAAGACATGAA	1309
Qy	1213	ATAGTGTGATGATGAGAGAACTTTACGATGTCGTTGTGTAAACCGAGGATGCTGCC	1272
Db	1310	TCTGTTTCAACTTAATCTTCGCGCAACATAATTTTGTATCTCTCGTCCAGAAGATTTAAA	1369
Qy	1273	GGAGATTTGGAGTTGGATTTGCTTTTGGCAAGGAGGAATTTGCAAGCTTTTGTATGATTAT	1332
Db	1370	GGAGCACTTGAGCTTGGCATTTCTTTGCTTGAAGATAGGATGAGCGCATTTTGCACCAT	1429
Qy	1333	TTTGGGAGGGAATAAAGGGT 1353	
Db	1430	TTTAGGAATTCATCAATAGT 1450	
RESULT 7			
AB026494 1622 bp mRNA PLN 04-APR-2000			
LOCUS AB026494 1622 bp mRNA for acyltransferase homolog, complete			
DEFINITION Gentiana triflora GAT106 mRNA for acyltransferase homolog, complete			
cds.			

AB026494	AB026494.1	GI:7415596	acyltransferase homolog.
KEYWORDS	Gentiana triflora	cdna	to mRNA.
SOURCE	Gentiana triflora		
ORGANISM	Gentiana triflora		
REFERENCE	1 (sites)		
AUTHORS	Yonekura-Sakakibara, K., Tanaka, Y., Fukuchi-Mizutani, M., Fujiwara, H., Fukui, Y., Toshihiko, A., Yamaguchi, M. and Kusumi, T.		
TITLE	Molecular cloning and biochemical characterization of hydroxycinnamoyl-CoA:anthocyanin 3-O-glucoside-6-O-hydroxycinnamoyltransferase from Perilla frutescens and diverse plant acyltransferase homologs		
JOURNAL	Unpublished (1999)		
REFERENCE	2 (bases 1 to 1622)		
AUTHORS	Tanaka, Y. and Yonekura-Sakakibara, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-APR-1999) to the DDBJ/EMBL/GenBank databases.		
	Yoshikazu Tanaka, Suntary Ltd., Institute for Fundamental Research; Wakayama-dai 1-1-1, Shimamoto, Osaka 618-8503, Japan (E-mail:Yoshikazu.Tanaka@suntary.co.jp, Tel:81-75-962-8807, Fax:81-75-962-8262)		
FEATURES	Location/Qualifiers		
source	1..1622		
	/organism="Gentiana triflora"		
gene	/db_xref="taxon:55190"		
	35..1474		
CDS	/gene="GAT106"		
	35..1474		
	/codon_start=1		
	/product="acyltransferase homolog"		
	/protein_id="BAA93452.1"		
	/db_xref="GI:7415597"		
	/translation="MAGNSEDIKVLEKCRVAPPPDAVEFTVPLSFDDMRWLISDAEH SLVLTAESSMDFDYLAGDHORDSYKFNDLIPOLPEPIVTSDEVLPLEALQVTFESN LHMFYRPHCPNSKFTISSKSLVLFKHLPLAGNLIPWDSDDRMPELRYKKGD TGICIRNLHVLGDSASFELHFNKLVLDKSNGLSKFLPLSLPNYDRSVVQDPFH IRKINERKLKLSQGTPTNPAINSAIKDEVRATFILPIDIMKLKFISSNNRLTGS SNKLSFTVTSALITCLSKSLDTVVREKVEDKHAANLCAFINCRQRAPPYIPQNY FGNCIVPCMGVSTHEQLVNEGLSVAATAIGDAIHKRLHDYEGILRGDWISPPRTSA APRSTLTVYVGSQNRVHDFDADFNGKLEKHESVSTNPNSATLILISRRRFRKGALEL GISLPKRMDFATIFTNFINSLHVRSL"		
BASE COUNT	478 a 345 c 322 g 477 t		
ORIGIN			
Query Match	10.1%; Score 153.8; DB 12; Length 1622;		
Best Local Similarity	49.4%; Pred. No. 1.2e-21;		
Matches	692; Conservative 0; Mismatches 652; Indels 57; Gaps 9;		
Qy	4	ACCACCCCTCTCGAATCCTCCGAGTGGCGCGCTCCAGCACGGTGGCTGAGCAGTCA	63
Db	56	ATCAAGTTCTTGTGAGAAATGCCGTGTTCGCCACACCGGACGCGGAGTTTACA	115
Qy	64	CTCCGCTCACCTTCTCGACATGACGTGCTGCATTTCCACCCCATGCTTCAGCTTCTC	123
Db	116	GTCCCACTGCTGTTTTCGACATGCGATGGTGTGATCTCTGATCGAGAACACATCTGCAT	175
Qy	124	TTCTACGAACCTCCCTGTTCACAAACCCGCCCTTCCTCGAAACCGCTTCGCCAAACTCAAA	183
Db	176	TTCTACAGATTCGCCCATCTTGTCCCAACTCTA---AATTATCATTTTCATCATTA	232
Qy	184	CAATCCCTTATCTTAACCCCTCAAAACACTTCTTCCCCCTTTTCATGCAATCTAATCTACCT	243
Db	233	TCGTCCCTTTTCCCTTGTCTCAACACTTTCTCCGTAGCCGGGAATTTGATTGGCCG	292
Qy	244	CTA---TCGCGGAGAAATCCCGGAGTCCGGTATCAGACGGTCTGACTCGGTTTCTTTC	300
Db	293	GTAGATTCCTCCGAGAAATCCCGGAGTTCGGTTTACAAGAAAGGGGACTCCGTTTCTTTA	352


```
SIFTENQTATLEMDYFRMSGSGSWIPWYLDVTVIAGSGRRHDLRLDHLPLVSNPKY
YDALNGVEILKMDNDPQNLGAPNDPLVSDPLIPNKRATPRIRKNKSHLIPITLAVVG
SLVLAHMFVGVGLVIMKKKSPNDSNMCPLPHGDDSTNTRPAKSLPADLCRRFSI
FEIKASNDFFDLKLIIGVGGFSGYKQIDGATLVAVKRLTISNQCAKEFELELEM
LSKURHLVSLIGCYIGDEDMVLVYEVMPHGTLDKHLFRDRTSDPPLSKWRLEIC
IGAAGQLXHTGAQYVYLIHRDKITNILLIDENETVKVSDGLSRVGTSAOPHVST
VYKTFGLDPEYVRQVRLTEKSDYSEFVVVLVLCRPIRQMOSVPEQADLHWVK
SNRRGTVDQIIDSLSADITSTSLERKFEIARVQDQRMERPMQSDVVMALFALQ
LHETAKKDNVGLDMLPMSEGEVGTGTTDGEDDLFSRTTGHVKGSTTDDSDSLVYVGDER
SGSWGVSEINEPKAR"
6701.. 7210
/notes="gene_id:MXF12.3
unknown protein"
/evvidence-not_experimental
/codon_start=1
/protein_id="BAB10825.1"
/db_xref="GI:10177546"
/translaton="MINLEKEAKEQIILRDGKRKIVHHVNTTLLVKIYEASAIIDRRH
LFLAPYGPNNYFIKAWLDDTKEYGYVVKDSFKVNPPLKIPLNEPAQYLYVELLK
GMSYRDPCTNRTVVMGRAKIRLSPLMTSYKERPISGIDVLGNSQCQVYKGYLTY
AMKLRHSGAD"
7670.. 10111
/notes="gene_id:MXF12.4"
/codon_start=1
/evvidence-not_experimental
/product="receptor protein kinase-like protein"
/protein_id="BAB10826.1"
/db_xref="GI:10177547"
/translaton="MNCNVLFLLSVLYTAGVTAAYHPTDVELFNCGDTSNVNDNSG
RNWTVESRQILSSNLVNASPTSEASYQAGVSRIPIYMKARIFRSEFTYFPVTPGSI
LRLYFTPTQYKSGFDVANSFYVSKVNGFTLLRNFNADSTVOASIPLSNLIKEFIIP
HQUNLFTFYKSNLLAFNNGIEIYVMPDRFYKSGFDNVLNRVSSDYDFQIDNSTAFE
SVHRLYVGQITNEVDSDGMRWMLSDSGNCSGIVNVPVKINYTEKTPASVAPYD
VYATSRMGNSNLNMFNTGMFLTDAGYVNLVRLHFCETLPQVTKAGORVESIFVED
KMAKETDVLRLSGDRIPMYLDPSVYVVGESGMIOPELRDLDLVPLKDTNQTYDAIL
SGVEILKINDSDGNLARNPELLVSTSDTSDSNVTPPIKGPVHLVILTVGVSGVG
LATIVEIIMLLIROMKRNKKNKNSVIMFLLKQYIYAEKLIKTKFSFHVTVGVGG
TVYRGNLSNRTAVVYKDLKNGDDFINEVTSQTSNHNIVSLFGFYEGSKRAI
ISEPLEHSGLDQFTSRNKLSTPNVTLYLGALGARGLEYLHYGCKTRVHFDIKPNQ
ILLDDNFCPKVADGLCKEKRESILSLDTRGTIGYIAPEVVSRYMGSHSKSDVY
SYGLMLDMDIAGARKVETTCNGSTAYFPDWIYKDLNGDQTNIGDEINEEDNKIVK
KMILYSLWCIKPCPSDRPPNKNVMEIEGSLDALELPPKPSRHISTELVLESSLSDG
QEAERQTLDSTII"
```

CDS

/notes="gene_id:MXF12.3

unknown protein"

/evvidence-not_experimental

/codon_start=1

/protein_id="BAB10825.1"

/db_xref="GI:10177546"

/translaton="MINLEKEAKEQIILRDGKRKIVHHVNTTLLVKIYEASAIIDRRH
LFLAPYGPNNYFIKAWLDDTKEYGYVVKDSFKVNPPLKIPLNEPAQYLYVELLK
GMSYRDPCTNRTVVMGRAKIRLSPLMTSYKERPISGIDVLGNSQCQVYKGYLTY
AMKLRHSGAD"

CDS

/notes="gene_id:MXF12.4"

/codon_start=1

/evvidence-not_experimental

/product="receptor protein kinase-like protein"

/protein_id="BAB10826.1"

/db_xref="GI:10177547"

```
translaton="MNCNVLFLLSVLYTAGVTAAYHPTDVELFNCGDTSNVNDNSG
RNWTVESRQILSSNLVNASPTSEASYQAGVSRIPIYMKARIFRSEFTYFPVTPGSI
LRLYFTPTQYKSGFDVANSFYVSKVNGFTLLRNFNADSTVOASIPLSNLIKEFIIP
HQUNLFTFYKSNLLAFNNGIEIYVMPDRFYKSGFDNVLNRVSSDYDFQIDNSTAFE
SVHRLYVGQITNEVDSDGMRWMLSDSGNCSGIVNVPVKINYTEKTPASVAPYD
VYATSRMGNSNLNMFNTGMFLTDAGYVNLVRLHFCETLPQVTKAGORVESIFVED
KMAKETDVLRLSGDRIPMYLDPSVYVVGESGMIOPELRDLDLVPLKDTNQTYDAIL
SGVEILKINDSDGNLARNPELLVSTSDTSDSNVTPPIKGPVHLVILTVGVSGVG
LATIVEIIMLLIROMKRNKKNKNSVIMFLLKQYIYAEKLIKTKFSFHVTVGVGG
TVYRGNLSNRTAVVYKDLKNGDDFINEVTSQTSNHNIVSLFGFYEGSKRAI
ISEPLEHSGLDQFTSRNKLSTPNVTLYLGALGARGLEYLHYGCKTRVHFDIKPNQ
ILLDDNFCPKVADGLCKEKRESILSLDTRGTIGYIAPEVVSRYMGSHSKSDVY
SYGLMLDMDIAGARKVETTCNGSTAYFPDWIYKDLNGDQTNIGDEINEEDNKIVK
KMILYSLWCIKPCPSDRPPNKNVMEIEGSLDALELPPKPSRHISTELVLESSLSDG
QEAERQTLDSTII"
```

CDS

/notes="gene_id:MXF12.5"

/codon_start=1

/evvidence-not_experimental

/product="receptor protein kinase-like protein"

/protein_id="BAB10827.1"

/db_xref="GI:10177548"

```
translaton="MICFILFVSFLVSVSATAPYKPDVFLINCGETDVPFONHGRF
WTOEKNILPKNSDNASFVSVYKESGIPQVPMTARIFRSDFTYFPVSPGWKFL
RLXPYPTYSKSGFDVANSFYVTVNDFTLLQNEFADITVKASIPESKSLIKEFTVPY
LTNLNTRPSNNSLAFVNGIEIYVMPDRFYKSGFDNVLNRVSSDYDFQIDNSTASST
VHRLVNGHVMDEVDNSGMRWMLSDDYELIGGVSPMPDVNISYTEKTPAYVAPAY
VYTCRMMGNAQDPTYNLNLNFWLTDAGFSYLVRLHFEKYLKNKANORVSEIFLG
NQAREMDVIRLGGPRPIYIDFRIYVGSSEGRPDRLDLHLVKNDRPEYFIAL
NGVBEILKNSNGLAIQDNELNPNPLSNLTPNHVTQOIKGKSHLLVKYIFIAVGP
GTGLATFVVLMLWROMKRNKREERVMEKLLNMYTAEKLIKTKFSFYIIGKGG
FGTVYGNLSNGRKYAVKVLKDLKGSADRFINEVASNSQTSNHNIVSLFGFEGRSRR
AIVYFLENSLDQFMGRNKSITQDVTLLYLGALGARGLEYLHYGCKTRVHFDIRP
QUNLLDNLCPKVSDFGLAKLCKEKRESVLSLMDTRGTIGYIAPEVVSRYMGSHSKSD
VYSMLVMDIAGARKSVETVETVDSASSTYFPDWIYKDLNGDQTNIGDEINEEDNKIVK
EIAKKMTVWGLWCIIQPCPSDRPSMNRVVENMEGSLDALEIIPKPSMHISTEVITES
LSGSGEDV"
```

CDS

```
join(16412, 16498, 16859, .16924, 17003, .17038, 17085, .17192,
17337, .17461, 17541, .17640, 17803, .17974, 18075, .18191,
18285, .18395, 18488, .18632, 18712, .18844, 18934, .19009,
19232, .19297, 19385, .19492, 19591, .19698, 19873, .20070,
20164, .20373)
```

/notes="gene_id:MXF12.6"

/codon_start=1

/evvidence-not_experimental

```
/product="ABC transporter-like protein"
/protein_id="BAB10828.1"
/db_xref="GI:10177549"
```

```
translaton="MGNKLIITGGSTKHGSSYRDPLQLNQBDKPKANGSENGLD
LEHGVBAANNGFGRVPAALCOYSSIMVSOAKPDAGKLVIGTIALIGSTTNLLVPKF
GMLTIDVSRDVKTPQOOTESLIAVRNAVITILLIVIGSICTALRAWLENSASERVV
ARLKDLFRHLHMQEIAFYDVTKGELLSRLSDETQIINKAATNLSALRNVTALLI
GLGPMFTSSWKLTLALLVWPVIVSAVQFGRYLRELSHTTQAAAASAAVTSFSGA
VTVYSGMTVGAITSFYLSLTGVSSVSSLSLTAMKAAGASRRVFIQIDRVSSMS
LTYTIGMTVGAITSFYLSLTGVSSVSSLSLTAMKAAGASRRVFIQIDRVSSMS
SSGDKCPVGNPDGDELNDVWFAPSPRSHMILKGISLRLTPGSKVALVPGSGGKTT
IANLERPYDPLKGIILLNGVLSMEISHQYLHKOISIVSQEPILFNCSEVENIAYGED
GEASFTDIENAKMANAHEFTAPDKYNTVYGERGLRSLGGQKQRIARAILATNPFS
VLLDEATSLDADSEYSLVQDAMDLSLMAGRTVLVIAHRLSTVTKADCAVVISDEVAE
KGTDELISLNGIYTNLVKROLQSSSSVTTI"
```

CDS

join(25348, .26359, 26401, .26756)

/notes="gene_id:MXF12.7"

/codon_start=1

/evvidence-not_experimental

/product="anthocyanin acyltransferase-like protein"

/protein_id="BAB10829.1"

/db_xref="GI:10177550"

```
translaton="WVNEEMESSLKVIDVARVTPSNSDSESLLPLTFFDLWYKLH
AVERVIFKLTDAERPEFDSVIVPNLKTSSSSUSHYLPAGKLVWPELPDPKPIVYT
PNDASFTVAESNADFSRLTKPEPPTTELYPLVPELHVSDDASASVQVTLFPNQG
FCISVNAHVALDGTNTNFKSWARTCKNQDSFLQDLPQDLIPVYDRTVIKDPMDLDTKI
LNAHVRKAVFTGKKEPENKSKLKLWSPGIDPVDVRYTLNLTREDIQKLERLKES
SSSSVSSSPKELRSTFVIVSYALTCLIKARGGDPSPVGYGFAPDCRSLMVPVPS
SYFETPNSGEFFLAARMVSDSVEALDENVALKIPEILEGFTTILSPGTQVLSVAGSTR
FVGVLDGFMGRPEKVVVVSDIOGEAISFAESRDGSGVGLFSLKHEMDVLVDLLH
KGLFN"
```

CDS

complement(join(27466, .27879, 28046, .29073))

/notes="gene_id:MXF12.8"

/codon_start=1

/pseudo

/evvidence-not_experimental

/product="transposase-like protein"

complement(join(30804, .30874, 30947, .31289))

/notes="contains similarity to Ac-like transposase"

gene_id:MXF12.9"

/codon_start=1

/pseudo

/evvidence-not_experimental

/product="anthocyanin acyltransferase-like protein"

/protein_id="BAB10830.1"

/db_xref="GI:10177551"

```
translaton="MVSLSNIEVARVTPSNSDASLSPLTYFDLIYKLRVERV
IFYRITWYTRPFDSVIVPNLKTSSLSCLSHYLPAGKLVWPELDHKTPTIVYSONDDV
SFSAVETNADFSSLSGNEPSPTELYPLVALOSSDDASIVSQVTLFPNQGFCIGV
SAHVALDGTNTNFKSWARTCKNQDSFLQDLPQDLIPYDRTVIKSPDSENKVLNEWR
SFTKLAGGKEPENVKSLNCPSEIIGDPVVVRYTLQTRDIOQLRLERLKRREYSSSS
STSSKLEPSTFVIVSYLVYICIRARGGEPRHPVGYAFSDVCSRLMNPPTNYFNG
CIAGCSMLMTAKMTMCEGGLAAATMVSDSIEWDESFAWKIPDFVAYATLPPEQTQ
ILVSGSNRFGVYELDFCGWRPDKVMVYVISPNGNISMAESRDQNGSVEIGFSLKHEM
DTLIDLHLCELT"
```

CDS

34363, .35709

Query Match

Best Local Similarity 7.0%; Score 106.4; DB 12; Length 66237;

Matches 664; Conservative 0; Mismatches 716; Indels 48; Gaps 8;

Qy 45 CACGGTGGCTGACGAGTCACCTCCCGCTCACCTTCTTCACATGACGTGGCTGCAATTTCCA 104

Db 34419 CTCGTCGGCGTCACTCACTCTCCCGCTCACCTTCTTCGCTCAACACAA 34478

Qy 105 CCCCATGCTTCAGCTTCCTCTTACGAACCTCCCTGTTTCCAAACCCGCTTCTCCTCGAAC 164

Db 34479 AGCGTCGAACGAGTCATCTTCTACAAACTCACCGACGTAATCGCTCTTTATTCGACTC 34538

JOURNAL

Submitted (15-JUL-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA

REFERENCE
AUTHORS

7 (bases 1 to 101176)
Cheuk, R., Shinn, P., Brooks, S., Buehler, E., Chao, Q.,
Johnson-Hopson, C., Khan, S., Kim, C., Altafi, H., Bel, B., Chin, C.,
Chlou, J., Choi, E., Conn, L., Conway, A., Gonzalez, A., Hansen, N.,
Howing, B., Koo, T., Lam, B., Lee, J., Lenz, C., Li, J., Liu, A., Liu, J.,
Liu, S., Mukharsky, N., Nguyen, M., Palm, C., Pham, P., Sakano, H.,
Schwartz, J., Southwick, A., Thaveri, A., Toriumi, M., Vaysberg, M.,
Yu, G., Davis, R., Federspiel, N., Theologis, A. and Ecker, J.

TITLE

Direct Submission

JOURNAL

Submitted (19-JUL-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA

COMMENT
FEATURES

On Jul 15, 2000 this sequence version replaced gi:2618677.
Location/Qualifiers

source

1. 101176
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"

/chromosome="1"

/clone="F21B7"

join(736..1038,1276..1713,1874..2119,2205..3032,

3117..3274,3483..3610,3725..4134,4313..4435)

/note="unknown protein; similar to ESTs dbj|AV529515.1,
dbj|AV529266.1, dbj|AV54950.1, dbj|AV523761.1,

dbj|AV523761.1, dbj|AI996703.1, and dbj|AV546743.1"

/codon_start=1

/evidence=not_experimental

/product="F21B7.1"

/protein_id="AAF86508.1"

/db_xref="GI:9280639"

/translation="MQPLPARSGDHEGFWNSHPLLVAAGDETNGTGLGHFSQNGSL
ARQSDSKAGDAINYPVTYRVSLSRSHSYVVLRFSSVCMIRCSRVAVGLANGI
YCDALRLNKFSLVTPVQVRQGTTRVNGTGPVAGPRLAYASKSMKTGR
LSDTSPSPSLSPSSSGSSGMARYAMESKOLANGLNGLMDGYKTLTSKCYQDML
PQSTSPASNAWKGVGSQDAENAGYAVKDLVSGALVSQFKAHTSISALCFDP
SGTLTAVTACGNNINVFQIMPGRSHNAPGDLSEWESHVLFKLHRTGITSATVQDI
CFSQSOQWVAIISKQCHIFVLNMSGDAAPCEGEETRLPASPSPWFTQSLSS
NQSLSPPTAVLSVSRIRKYSFGWLTNTVSNATATGKVFVPSGVAVAEHSVTH
DLQNSRTNALEHLVTPSGHVQVHELLSVCTESPENGLRVQKTSHVQVQDDLRV
KVEPIQWDRQWLETERLPKSLTEKOYDLETVTNSHLTSHEDACLSDMNSHFS
EDKYLKSCSKPERSHCYLSNFVETVSGMLPWQNSKISFHVMDSPRDSSTGGF
EIEKPAHELEIKQLLPVDFHSTKATLEDRFSMKCYHTSATGSHOVNGKTCQDI
INCHSKPGSTESAESSPEGSTKQENLHSDHMSNSISLPLVPTVNGIYKTEKN
ANGMEKPVAKLSTLKTETITNGFTTPPILITDSVNEQMLSTGKPPMGFGFALHEHC
KAVADPKEEHLKKLDEVTNVHLNVNNNTEKLGDKMVSQVNLNAGVCILYKDFL
RVGNFNSVFPLLYKTLWRK"

complement(4838..6223)

/note="hypothetical protein"

/codon_start=1

/evidence=not_experimental

/product="F21B7.2"

/protein_id="AAF86541.1"

/db_xref="GI:9280672"

/translation="MASCIDELHFTHLHIPVTINQQLVHPSPPTANQSPHLSLYLS
NLDDIICARVFTSVYPTNNRESFVLKRLQDALSEVLVVPYPLSGRLREVENGLK
EVFPEQGVLMVYANSMDLADLITVPNAPLPLIFNPGEEAYKILEMPLLIQA
VTFETCGFSLGIRLCHICIDGPGAMQFLGSAWAKTAKGLIADPEVPWRETFKPN
PPWKYPHHYLPTEERSNLNMDTKPLQKCYRISKEFCQKRVCSAQGEDPLVCS
TFDAAAHIRWSKALDKPLDYNLRKLTFSVNRTRLETKLRKKGYNVVCACAM
SVSEFLNLSLTKTRLVQDARLRSDEYLRSDVYDVRKPRLEFGGKLITQWTR
FENYETADFGWGKPVVAGIDLRPTQVCVLLPQGGVESGNDQSMVVCICLPPTAVHT
FTRLISLDHK"

complement(join(6700..7357,7482..7884,8620..8839,

8920..9241,9758..10380))

/note="unknown protein; similar to ESTs gb|T04168.1,

gb|T21964.1, and gb|AW004462.1"

/codon_start=1

/evidence=not_experimental

/product="F21B7.3"

/protein_id="AAF86540.1"

/db_xref="GI:9280671"

/translation="MGHDSFCYLIVLRCAIRCGIIALMQICALQKQKRRSKMESDRS
SOAKAFETKTVGLVAGSITKEIPAMFHTPDTLSTLKQTAPPSOULTIPTVDLKG
SMDLISRSVTEKIGDAERWGFQVNVHGISVEVMEKREIRRFHQDEPEVKREY
SRDTRDLYVSNIDLTCHKANWRDTLACYMAPDPKLODLPAVCGEIMMEYSKQL
MTGLEFLELLSEALGNHNLKDMGCAKSHIMFGQYPPCPQDPLTLGISTGDFDFSF
ITILLQDNGIGLQVHDQCVWVPGVALVINDLLQILISNDKFSIAEHRVITANGS
SEPRISMPCFVSTPMKPNRIYGPICKELSEONPAKYRDLTITEFSNTFRSQTIRTE
IPALFRATPATLASLSPPPKHILTPTVDLKGASVVEKIGEAEEKGLFHLVHNP
VEVLERMQIRGHEQEPKAKRFYSRDTRDLYVFNHDLQNSEASWRDITLGCYT
VGPRLLEDPAVCGEIMMEYSKREIMSGELFELLSEALGNHNLKDMGCAKSOQM
VGOHYPPQPDPLTIGINKHTDTSFLTLQDNGVGLQVHEQYHIDVTVPVGLVIN
IGDFQKIKMLLRSCFISCMIMGLILFYCVQLITNDKFSIAEHRVITANGSEPRK
SVAIVSTFEMRAYSRVYGPICKLLESAENPAKYRDLTITEFSNTFRSQTLPKLUHFK
I"

10833..11348

/note="unknown protein; similar to EST emb|Z30724.1"

/codon_start=1

/evidence=not_experimental

/product="F21B7.4"

/protein_id="AAF86509.1"

/db_xref="GI:9280640"

/translation="MRGFSVSSHLWTSVEIGEVFLSGTTLWDAVAEADVDMWVES
ACLMVKNKDLRLVPAEGLSEPSFFSLRFLSPSIRIEGCFAGKFSDLRFRWE
AWKCVVLRGWLQDQIGRWSVGNWVSVRFVAGPASSLVRLDRQDSYTSANDLVVVV
LQSSGPGIRRY"

join(111965..12033,12173..12303,12448..12519,12596..12685,
12756..12915,12985..13041)

/note="similar to ATHP1 dbj|BAA37110.1"

/codon_start=1

/evidence=not_experimental

/product="F21B7.5"

/protein_id="AAF86510.1"

/db_xref="GI:9280641"

/translation="MNTIVVAQLQRQDYIVLSYQOGLNDQPSSELRKLQDEGTDPF
VAEVSLFFDDCSKLINTMSISLERPDVDFQVDSGHQLKGSSSSLCFLVSGARR
VKNVCSIFKCCDDQVNEGYSNHTNDNDINHRCLRCLQQVDYKMLKTKLQDLF
NVSPKLCENLVLLEKQILQAGGTIPQVDIN"

16495..17682

/note="similar to protein kinase pir|JAS57676; similar to

ESTs gb|AI993651.1, dbj|AV538995.1, dbj|AV522524.1, and

dbj|AV538995.1"

/codon_start=1

/evidence=not_experimental

/product="F21B7.6"

/protein_id="AAF86511.1"

/db_xref="GI:9280642"

/translation="MRTKLWCLMFLRLRGFTTEALDPVDFLALQAIRKSLDDLPG
SKFFESWDFSDPGFAGVYNGDKVLSLNLGDDPRAGSLSGRIDPAIKLSALTET
SIVPGRIMGALPATISQKDLRLAISRNFIISGEIPASLGSEVRGLRTLDLSYNLQST
ISPSIGLPELSNLIILCHNLTGSIPLPQLTQTLRIDLRKNSLTGSIASLPPSLQY
LSLAWNLTGVSVHVLRLNQLNYLDLSLNRTGTIPARVFAPITNLQLRNPFGL
IQPANOVTISTVDLSYNRFSGISPLSSVENLYLSNRFETGEVPASFVERLLSANIQ
TLYLQHNFLGIGIISPAAEIPVSSSLQVNCMVPPLOTTCPLKAGPQKTRPTTQCTE
WRG"

complement(join(18462..18515,18612..18742,18840..18975,
19064..19139,19214..19410,19497..19595,19668..19752,
19865..20103,20179..20253,20345..20649,20798..20902,
21088..21216,21356..21437,21513..21608,21780..21849,
21929..22025,22408..22502,22586..22659,22810..22888,
23041..23101,23206..23315,23411..23556))

/note="hypothetical protein"

/codon_start=1

/evidence=not_experimental

/product="F21B7.7"

/protein_id="AAF86539.1"

/db_xref="GI:9280670"

/translation="MPEKFKRRSLNFICLPKLKECARGOILGYKMETKAI GFHLSLP
PQNVICGHTLTAVFVNNHQILILFGGSTAVANNHSLPEISLDGVNSVHVSFVLTR
KWRNLPIGDVPSPRACHAALYGTLLIOGIGPGSPGDVYMLDMTNKNKFKLV
GGTTPSPRYGHVMDIAQORVLVFSGNGNEILDWALDTRGPFSDRLNPNQGPS
GRMYASSSREDGIFLCCGIDHSGVTLDGTYGLKMSDSNDVWTPVPAVAPSPRQHTA
VFGSKLHVIGGLNRLARLDGEAVADTNQPETSSAGSANKRQNOYQLMRKCHHAASF

CDS

CDS

CDS

CDS

Query Match 6.7%; Score 102.2; DB 12; Length 101176;
Best Local Similarity 45.4%; Pred. No. 7.5e-11;
Matches 620; Conservative 0; Mismatches 718; Indels 27; Gaps 6;

QY 7 ACCCTCTCGAATCTCCGGAGTGGCGGCGCTCCAGGCACGGTGGCTGAGCAGTCACATC 66
DB 37271 AACATCATCGAGACCTGTGCACATCTCTCCCAAGGCGACCGTTCGATCCACCACTCTT 37330
QY 67 CGGCTCACCTCTCGACATGATGGCTGCAATTTCCACCCCATGCTTCAGCTTCTCTTC 126
DB 37331 CTTCTCACCTCTCTGATGCCCCCTGGCTCTCTCCCACTCGCGGATCTCTCTTCCT 37390
QY 127 TACGAACCTCCCTGTTCCTCAAAACCGCTTCCTCGAAACCGTTCCTCCGAAACTCAAAACAA 186
DB 37391 TTCTCTTACCAAACTCAACTGAATCTTTCCCTCCAGACTTTGTACCCCAACCTCAACAT 37450
QY 187 TCGTTATCTTAACCTCAACACTCTTCCGCCCTTTTCATGCAATCTAATCTACCGCTCTA 246
DB 37451 TCCCTCTCCATCACTCTCCAACTTTCTTCCCTTACGCGGTTAAACTGATTTATCCCGCT 37510
QY 247 TCGCCGGAGAAAATGCCGGAGTTCGGGTATCAGAACGGTGAAGTCTCTCTTTTTCAGGAT 306
DB 37511 CGTCCGGACCTCCATATTTACACTACAACGCTGGGAAGACTCTCTCTTTTTCACCGTA 37570
QY 307 ATGGAGTCT-----GTCGGAGATCATCCGCTATCCGCTCATAAATPACTACTGCTTT 357
DB 37571 GCAGAGTCTACTGAAACGACCTTCGACCAACTGAAGTCTGATTCCTTAAAGTATACAT 37630
QY 358 GCCCTAGGACGATATGAAGATCTCCAGCTGCGCGCGGATAGTGAGGAACTCTGATCGG 417
DB 37631 GTGTTCATGCGCTCTTGCCCAAGTTACCTCTCTCAGCTCTCTCCGGAAGAAATTCAG 37690
QY 418 AATTTGTTTCAAGTTTACCGCTGCAAGTACTCTGTTTCCCGGTGCGGGGTGTCATC 477
DB 37691 A---TGCGACCAATATGGCAATGAAGTACCACTCTTCCCTGGAGCTGGAATCTGTATA 37747
QY 478 GGAATTAACGACGACACACACCGCTAGCGATGCTCCATCTGTTGAGGTTTATGAAGACT 537
DB 37748 GCGAAGTCTACATCTTTGACGAGATGAGTCACTTCCAGTCTTTCATGAAGTAT 37807
QY 538 TGGGCTTCATCACTAAATTCGGAGAGATGATGAATTTCTTGGAGGAAAGGTGAATCT 597
DB 37808 TGGATGTCTATGACCAATTCAGCGGCAAGATCGCGCCACGCTTCTTCTACCTCTCTG 37867
QY 598 TTGCGCGTTTTCGACCGTCTGCTGTAATTTCCGCTTAATTCGACACATATTTATCG 657
DB 37868 CTTATTCAGCTGCGAGAAACATATCAAGGACCCAGGAGAGGTAGCCCGACGACATTTG 37927
QY 658 AACACGCGCAGAAACGTCGTTGGAATCGCAGCATCCATCTTTACCGACGATCGGAT 717
DB 37928 GAGCGTTTGGAGCCAAACTCTGCAAAACACAGTTTCACATGTTACGCTGAGAACATG 37987
QY 718 C---GAGCTACTACCTTTTACCCCAATCTGAATTTAAGAAATTTGAAGGTTTGTATCAG 774
DB 37988 GTACAGAGCCATTTTACATTTAGCGCGGAAGCAGATAGACAATCTCAAGAGTTGGTTACA 38047
QY 775 AGAAAGCCCCAAATAGTATCTCTTCTTCTGCTGCGATCGCAGCTTATATCTCG 834
DB 38048 GAGCAGCTTGAGAACTCAATCTCTGTTTCTACCTGCTGCTGCTGCTGCTGCTGCTGCT 38107
QY 835 ACCGCGATCGCAAAATCGCTGCGAGATTAACA--AAGACGCTGATGACACAAACGCGT 891
DB 38108 GTAAGCTTGATTAAGACACTTTTCAAGACAGTGAACAGAGCGCAAGGCAAGTCTTT 38167
QY 892 TTCTTTTAAATTCGATCATTTAAGCGCGCTTTTGA-----TCCGCGCGCTCGGGG 945
DB 38168 CACTTGATGATCAACGCTGATTCGAGGAATCTCTCAAGTACACACAAACCAATACCACAA 38227
QY 946 AACTACTTCGAAACTGTCTATCTGTTTGGGATGGCGGAAGATCCTGCGGCGGATTTGGTC 1005
DB 38228 ACATACTTTGGCAACTGTATGGCTCTCTGATCTGATCATGATCAAGAAACACGATTTGCTA 38287
QY 1006 GGAGATGAAGGGTGTCTTTCGGGCGACTGAGCGGATCGCGCGGCAATAGAGAGACG 1065

DB 38288 GGAGAAAATCGTCATGCTCTCTCAGATGCAATCACAGCGAGATCAAGATATGTTA 38347
QY 1066 ACGCACAGAAGATCTCTAGAACTGTGGAGAACTGGCCCTCTGAGATTCGCGAAGCCTTG 1125
DB 38348 TCAAGCGATCTCTTGAAGACAGACCAAGATGGGACAAAGGAGTAGTAAATGGGTCTATG 38407
QY 1126 CAAAACCTGTATTTCTCGTGGCGGATCGACAGGCTTTGATCTTTACGCGCGGATTTT 1185
DB 38408 TCCATTATCCAACTCAJTTGAGGAGCTCCGAAATTCGGACTGTATCAGATGGATTTT 38467
QY 1186 GGATGGGTAAAGCGGTGAAGCAAGAGATCTCTCGATTTGATGGAGAGAAATTTACGATG 1245
DB 38468 GGATTAGGAAAGCCCTGCAAAATGGAGATAGTGCACATAGAAACAG--GTGGTCTCTATC 38524
QY 1246 TCGTTGTGTAAACCGAGGATGCTCCCGAGGATTTGGAGGTTGGATTTGTTGCCAAAG 1305
DB 38525 GCATTTCTCAGATCTCCAGACGCGCAATGAGTTGAGATTTGGAATGACATAGAGAAG 38584
QY 1306 GAGGAATTCGAAGCTTTTCATGATTTTTCGCGAGGGAATAAG 1350
DB 38585 AAGAAATGATGTGTTGACTCACTTCTACAAAGGGATCAAG 38629

RESULT 10
AC003027
LOCUS AC003027 119914 bp DNA PLN 30-JAN-1999
DEFINITION Arabidopsis thaliana chromosome I BAC F21M11 genomic sequence,
complete sequence.
ACCESSION AC003027
VERSION AC003027.1 GI:4079614
KEYWORDS HTG
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 119914)
AUTHORS Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altai, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vysotskaia, V.,
Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 119914)
AUTHORS Federspiel, N.A., Palm, C.J., Conway, A.B., Kurtz, D.B., Conway, A.R.,
Au, M., Araujo, R., Buehler, E., Dewar, K., Feng, J., Kim, C., Li, Y.,
Oji, O., Osborne, B., Shinn, P., Sun, H., Toriumi, M., Vysotskaia, V.,
Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1997) Biochemistry, Stanford University/DNA
Sequencing and Technology Center, 855 California Avenue, Palo Alto,
CA 94304, USA
REFERENCE 3 (bases 1 to 119914)
AUTHORS Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altai, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vysotskaia, V.,
Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1998) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
REFERENCE 4 (bases 1 to 119914)
AUTHORS Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F.,
Altai, H., Araujo, R., Huizar, L., Rowley, D., Buehler, E., Dunn, P.,
Gonzalez, A., Kremenetskaia, I., Kim, C., Lenz, C., Li, J., Liu, S.,
Luros, S., Schwartz, J., Shinn, P., Toriumi, M., Vysotskaia, V.,
Walker, M., Yu, G., Ecker, J., Theologis, A. and Davis, R.W.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

COMMENT

On Dec 30, 1998 this sequence version replaced gi:2734094.
 Bases 1-9262 of clone F21M11 overlap with bases 68998-78259 of
 'TAMU' BAC clone F20D22 (AC002411) and bases 119525-119914 of clone
 F21M11 overlap with bases 1-389 of 'TAMU' BAC clone F21B7
 (AC002560).

e-mail for correspondence: arabidopsis.stanford.edu
 Genes with similarity to proteins in the databases are described as
 'putative', '-like' or 'similar to'. Genes that have EST
 similarity but no significant protein similarity are described as
 'unknown proteins'. Genes that are annotated based only on gene
 prediction software are described as 'hypothetical proteins'.
 The software programs used to predict genes include: Graal
 (informatics group, Oak Ridge National Laboratory,
<http://compbio.ornl.gov/section/index.html>), GENSCAN (Chris Burge,
<http://compbio.stanford.edu/~chris/GENSCANW.html>), Fexa (V. Solovyev
 & A. Salamov, Sanger Centre, <http://genomic.sanger.ac.uk/>), and
 NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, <http://www.cbs.dtu.dk/NetPlantGene.html>).

FEATURES

Location/Qualifiers

1..119914
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="F21M11"

misc_feature

1..7000
 /note="Proteins in this region are annotated in the F21M11
 entry, AC002411."

misc_feature

1..9262
 /note="Overlap with bases 68998-78259 of 'IGF' BAC clone
 F21M11."

gene

7684..10913

/gene="F21M11.1"
 join(7684..7866,7965..8027,8122..8160,8250..8381,
 8458..8728,8875..9893,10003..10256,10344..10410,
 10501..10551,10685..10781,10885..10913)

CDS

/gene="F21M11.1"

/note="Hypothetical protein"

/codon_start=1

/protein_id="A010665.1"

/db_xref="GI:4204284"

/translation="MAPNLRKACDAMKLLGSETKTRAFRLKLLKTYENNNDFTIE
 DAYKVLDAJFLADQATEKNEKKEKRSVATSGRRKAPLEPLQDRE
 DMDDEDFPLRLRRGRASSSSSSNNEDLKTQPEEDEDGVTLEPLKRY
 VRRNGRLATVYNNAPSSSRLSMEPEVPMWLLPAHPWTKVSEASALVIND
 ENRIDHLPVTSIDVGCNAPMLEKSNHIVQEDWETKDLINDTANDVSPSSAIGS
 SEHKAASVELASSTGEAKICLSFAPATGETTNLHPLSMEDLRRAMEKCLSKYKI
 VHPFVSLGPMKDMCSYIDLAKNSTQLLETETVCDMSKAGDESGAVGSMPLVVP
 ECEISDGWKAISNMKIDITAGEENVEIPWYNEINEKVPFRYMPHSFVQDAPVIFS
 LSSFDQSCSTCIEDCLASEMCAIGVNGFAVTLGLLKEEFLEARISEARDO
 RKQVLRCECPLERAKKVEILEPKGLKARGAKCEKFCGCTKCGNRVVOGMHN
 KLOVFTPNKGWGLRTLEKLPKGAFCIEYIGELITIPELYQSFEDKPLPVLDAH
 WDSERLEGKALCLDGMFYGNISRFLNHRCLDANLIEIPVQETPDQHYHLAFTT
 RGTEAMLEAWDYGIDFNNDLSMKPDCICGSRFCRKRKRSTKTMILNKA"
 join(13033..13488,13717..13926,14139..14514)
 /gene="F21M11.2"

mRNA

13033..14514

/gene="F21M11.2"

join(13061..13488,13717..13926,14139..14316)

/db_xref="GI:4204285"

/gene="F21M11.2"

/note="Similar to acid phosphatase; Location of ESTs

110C2T7, gb|T42036, and 110C2XP, gb|A1100245"

/codon_start=1

/protein_id="A010666.1"

/db_xref="GI:4204285"

/translation="MDRTMELSLTIASLLGVVSGADWNILNQLGLGSSSQNGIYS
 KGKTDLKGCESEINVEINIRKFFVDPQECVSHIKDYMTSQYKDDVAVWDEVI
 LHTGSMCCSKGCDGMDAIFDIDDTLLSTIPYHKNGFGGKLNSTKVEDYDOKKK
 APAPVHKLYHDIRERGIKFLISSKEVLRNTVDNLIOAGYGNLMLRGLEDO
 QKEVQYKSEKRWLSLGRVWGMDQWSSGCPPLPRTFALPNSIYIYA"
 complement(join(14004..14312,14565..14641,14787..15831,
 15906..16006,16327..16362))
 /gene="F21M11.3"

mRNA

gene

complement(14004..16362)

/gene="F21M11.3"

CDS

complement(join(14567..14641,14787..15831,15906..16006,
 16327..16362))

/gene="F21M11.3"

/note="Unknown protein; Location of ESTs 40C3TT,

gb|AA728590 and 40C3T7, gb|T04573"

/codon_start=1

/protein_id="A010667.1"

/db_xref="GI:4204286"

/translation="MLDQISGHTHEKLSVETAPHLNDRVHIVEEIPKASVPIPE
 ICDEAEKCSPTISRRKRVFDSKVTEYHVVSSELSSEKNEEVEKRSKSSK
 TDQIIEVASNSGSGYENHRYKNCRESDDIEDRDCSDLDDEEYYSVDFGSE
 DSLHNPKEVYTDIGDQTEIDSKLRSSNETVRDGNHYDGGVLPNPENLTOKWSAK
 SKGTGKQKQKENSFIADQEKRRDSSFGTDQIDDIPLSVKPCRIEPLKLRNOE
 LAVDASUTWLSTSESGECSASMYLTITPEKLKSTSCYSLKRLNHDDRPVLCALIL
 IEDKQFSATSPKRSKSPDPTIIGTVGGYWGNSKADICGSASSFKGIPNTSSKY
 REDKSNVHSTPFPEARLEKALNNIDK"
 complement(16835..20238)

/gene="F21M11.4"

CDS

complement(join(16835..17185,17274..17392,17491..17656,
 17795..17885,17982..18079,18175..18361,18504..18604,
 18705..19049,19134..19349,19439..19791,19862..19967,
 20188..20238))

/gene="F21M11.4"

/note="Hypothetical protein"

/codon_start=1

/protein_id="A010669.1"

/db_xref="GI:4204288"

/translation="MGSSSPPEARAAQVPSMILFLEICTVHTVYNNRRLNDRVLSA
 NLNPKVTRRDLRFMESVTSYKSLAAVSVHLPOIQIPNDCYNKDALNNSPKH
 GESEDSMTDKVSKRGSDTSSRDCGSPLETSRDPKHODMTKQLSDHLLYE
 FESEYDAAHTPESYTEQAANKVRDITASQPSNAARKRICGDSFIESSEPNKTQDP
 TLRLMESLRSDPTDIVKAOONHMYFSEFWIYSFVLCRLNRLCVIIFVQQLPK
 SHTEQSKRRDITASDAMENHLKVPKRENNLMQSDIDCNGKCSANSDDLSEKTS
 KALEQTSNITICGFCASRVSEGLHYSGRPVDDGDIIFRSNVIHVSACIEWA
 POVYEGDTVKNLKAELGMKIKCTCKSLKGAALGCVFKSCRRSYHVPARETSRCR
 WDYEDFLLCPAHSSVFPNEKSGHRSRAEPLKINPAELCSLEQTPATKELVLCG
 SALSCKSKLMESLAVRFNATISRYWNPSYTHVASTDEKACRTLKLVLGILNGK
 IIRNAHMKSLKASQPVDEEPEFIQIDTQCGQGPRTARLARAKTKPLFGLAFYFF
 GEVYTIQWRANDAEALASQTSRLVGTWVLESYAGYKLLHPVIG"
 20752..24647

/gene="F21M11.5"

CDS

join(20752..20994,21362..21497,21596..21740,21825..21994,
 22102..22178,22542..22758,22920..23064,23343..23442,
 23599..23693,24021..24072,24227..24298,24441..24647)

/gene="F21M11.5"

/note="Hypothetical protein"

/codon_start=1

/protein_id="A010668.1"

/db_xref="GI:4204287"

/translation="MGANSKSVTASFVIAVFFLICGRTAVEDETEFHGDYKLSGI
 IIFGFASTQLRAWSILDCPYTPLDFNPLDLVLDITTKLLSANCWFCMVDLPYNOTD
 HPEKSRPDSGLSAITELDPGYTLPSTVWKWKLKCVFEGTEANAIVAVPYDMRLS
 PTKLEERDLYFHKLKLTFETALKLGRGSPSIVFAHSMGNVFRYLEWLRLEIAPKHYL
 KWLDOHIAFYAVGAPLLGSVEAIKSTLSGVTGELPVEGSTARLLSNSPASSLWMPF
 SKNCKGNTFTWTFSGGAARKDRVYDEEYQSKSYSGMPTNLINIEIPSTARELA
 DGTFLKAIEDYDPDSKRLHQLKKYVFFVIRNIAHRSLSAGLLELHYDDVPFNPLTW
 ERPPKKNVFCIYGAHLKTEVGYFAPSGKPYPDNWIITDIITYETEGSLVSRSGTVDG
 NAGPITGDETVPYVHSLKSNWLCGPKVNITMAPQILGKIQKQOEHGSDVHVELNVD
 HEHGSJLIANNKAPRVKITYTFEEDSESIPGKRTAVWELDKSGY"

27777..28734

/gene="F21M11.6"

27777..28734

/gene="F21M11.6"

28007..28465

/gene="F21M11.6"

/note="Unknown protein; Location of ESTs 203I24T7,

gb|H76794 and 203I24XP, gb|AA605510"

/codon_start=1

/protein_id="A010670.1"

/db_xref="GI:4204289"

mRNA

gene

CDS

/translation="MNTKMRLLPPRRVLITADKKRERDAFISVVDNPPPIAKFPSPPP
KLPPVPNPISKSSATAAEPIGSNQLMLAGYLSHEYLTQGTLFGEQNRQARAQESS
KIKPSHTVPEAECEPKRKRYREVANLLRSQAQLPGIVNPAQLAREFLK"
complement(29264..32033)
/gene="F2L11.7"
complement(join(29264..31015,31312..31414,31484..31587,
31782..32033))

Query Match 5.8%; Score 87.8; DB 12; Length 119914;
Best Local Similarity 44.9%; Pred. No. 7e-08;
Matches 617; Conservative 0; Mismatches 717; Indels 39; Gaps 6;
QY 11 TCCTCGAATCTCCCGAGTGGCGCCGCTCCAGCGACGCTGGCTGAGCAGTCACTCCCGC 70
DB 45957 TCATCGAGACCTGTACATCTCTCCCAAGGCGACCGTCCATCAACCACTCTTCCTC 46016
QY 71 TCACCTCTTCGACATGAGCTGGCTGCATTTCCACCCCATGCTTCAGCTTCTCTCTACG 130
DB 46017 TCACCTTCTTCGATGCCCTGGCTCTCTCCCACTCCCGATTCCTCTCTCTCTCT 46076
QY 131 AACTCCCTGTTCCAAAACCGCTTCTTCGAAACGCTGCTCCGAACTCAAACTCAATCT 190
DB 46077 CTATCCAAACTCACTGAATCTTCTCTCCAGACTTTGTACCAACCTCAAACTTCCC 46136
QY 191 TATCTCTAACCTCAAACTCTTCTCCCTCTTTCATGCAATCTAATCTACCTCTCTATCGC 250
DB 46137 TCTCCATCACTCTCCAACATTTCTCCCTTACGCGGTAACTGATTAATCCGCTCGTC 46196
QY 251 CGGAGAAATGCGGAGTTCCGGTATCAGAACGGTGACTCGGTTCTTTCACGATTAATGG 310
DB 46197 CGGACCTCCATATTTACACTACAACGATGGCAAGACTCTCTTGTTCACCGTAGCAG 46256
QY 311 AGTCT-----GTCGGAGATCATCCGATTCGCTCATAAATACTACTGCTTTGCC 361
DB 46257 AGTCTACTGAACCGACTTCGACCAACTGAAGTCTGATTCCTTAAAGATATCAGTGTG 46316
QY 362 CTAGCGACGATTAAGATCTCCAGCTCCGCGGATAGTCGAGGAATCTGATCGGAAT 421
DB 46317 TGCATGGCTCTTGCCCAAGTTACCTCTCTCACTCTCTCCGGAAGAAATTCAGA--- 46373
QY 422 TGTTCGAAGTTTAGCGGTGCAAGTCACTCTGTTTCCCGTCCGCGGTGTCATCGGA 481
DB 46374 TCGGACCAATAATGGTATGCAAGTCAACCATCTTCCCTGGAGCTGGAATCTGTATAGGA 46433
QY 482 TAACGACGACACACCGTTAGCGGTGCTCCATCTGTTAGGTTTATGAAGAGTTGG 541
DB 46434 ACTCAGCTACACATGTTGTAGCAGATGGAGTCACTTCACTATTCATGAAGTATGGA 46493
QY 542 CTTCCATCACTAAATTCGGAGGAGATGATGAATCTTTGGACGGAAGGTGAATGTTGC 601
DB 46494 TGTCAATTGACCAATCCAGCGGTAAAGATCCGCGCACGTTCTTCTACCTCTCTGCCA 46553
QY 602 CGGTTTTCACCGATCGCTCGTAATATTCGGCTAAATTTGGACACATATTTATGGAACA 661
DB 46554 TTCACAGCTGCAAGAACATGATCAAGGACCCAGCGGAGGTAGCGGACGACATTTGGAGC 46613
QY 662 ACGCGCAGAAACGTCGGTGGAAATCCAGATCCATCTTTTACCAGCGGATCGGATTC--- 718
DB 46614 GGTTTTGGAGCCAAAATCTGCAAAACACAGTTCACATGTTAGCCCTGAGAATCGTCA 46673
QY 719 GAGCTACCTACCTTTTTCACCAATCTGAATTAAGAAATTAAGAGGTTTGAATTCAGAGAA 778
DB 46674 GAGCCACTTTTACATTTGAGCGGGAAGCAGATAGATAATCTAAAGAGTTGGTTACAGAGC 46733
QY 779 AAGCCCAATATGATGTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 838
DB 46734 AGTCTGAGAAATCAATCT 46793
QY 839 GCATCCCAATCCGTCGGAGTTA-----CAAGACGCTGGATGACGACA 883
DB 46794 GCTTGATCAAGACACTGTACAGACAGTGAACAAAGGCCAACGAGGAACAGAGATG 46853
QY 884 AACGCCCTTCTTTTAAATTCGATCGATTTAAGGCCGCGGTTTGGAT-----CCGCCG 937

DB 46854 AAGTATTTTCACTTGATGATCAATGTGGATTGCAGGAATCGCCTCAAGTACACACACCAA 46913
QY 938 CTCGGGGAACTACTTTCGAAACTGTCTATCGTTTCGATGGCGAAGATCCTTCGCGCGG 997
DB 46914 TACCACAAACATACTTTTGCCTCACTGTATGCTGCTCGTATCGTATCAGTCAAGAAACAG 46973
QY 998 ATTGTGTCGAGATGAAGGGTGTTCGGGCGAGCTGAGCGATCGCGGGGGAATAGAGA 1057
DB 46974 ATTGTCTAGAGAAATCGTTTTCGGCGCTTCAGATCAATCAACGAGAGATCAAAG 47033
QY 1058 AGAGGACGAGCAGAGAAGATCTTAGAACTGTGGAGAACTGGCCGCTGTGAGATTCGG 1117
DB 47034 ATATGTTATCAAGCGATCTGTTGAAGACAGCACCAAGATGGGACAAAGTAGCTAAAT 47093
QY 1118 AAGCCTTCGAAACTGTTATTTCTCGTGGCGGATCGAGCGCTTGTATCTTTACGGG 1177
DB 47094 GGTCTATGTCCTTATCCAACTCAATTCGAGGAGCTCCGAAATTTGGACTGTATGACA 47153
QY 1178 CGATTTTTCGATGGGTAAAGGGTGAAGCAGAGATCTGTCTGATTTGATGGAGAGAAGT 1237
DB 47154 TGGATTTTGGATTAGGAAGCCCTGCATAATGGAGATCGTCACATAGAAACAG---GTG 47210
QY 1238 TTACGATGCTGTTGTATAACGAGGATGCTGCGGAGAGTTGGAGTTGGATTTGCTTT 1297
DB 47211 GTTCTATCGATCTTCAGAGTCCAGAGACGGCAATGGAGTTGAGATTGGAATAGCAC 47270
QY 1298 TGCCAAAGGAGGATTCGCAAGCTTTTGTGATGATTTTTCGGGAGGGAATAAG 1350
DB 47271 TAGAGAAAGAAATGGATGTTGTTGACTCAATTTACAAACAGGGGATCAAG 47323
RESULT 11
AB028618
LOCUS AB028618 85690 bp DNA PLN 27-DEC-2000
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 3, Pl clone:MOD1.
ACCESSION AB028618 BA000014
VERSION AB028618.1 GI:5041971
KEYWORDS Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
SOURCE clone:MOD1.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (sites)
AUTHORS Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S.
TITLE Structural analysis of Arabidopsis thaliana chromosome 3. II.
JOURNAL Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
MEDLINE and BAC clones
REFERENCE DNA Res. 7 (3), 217-221 (2000)
20363099
REFERENCE 2 (bases 1 to 85690)
AUTHORS Kaneko,T., Kato,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1999) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MOD1
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi).

Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is T1310 and the 3' clone is T26612.

FEATURES

source

```

1. 85690
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MOD1"
/clone_1lb="Mitsui p1"
2123. 8637
/note="contains similarity to TNP2-like transposon
protein
gene_id:MOD1.1"
/codon_start=1
/pseudo
/evidence-not_experimental
join(10671..10844,10921..11208)
/note="contains similarity to unknown protein
emb|CAA16722.1
gene_id:MOD1.2"
/codon_start=1
/pseudo
/evidence-not_experimental
15142..16497
/note="gene_id:MOD1.3"
/codon_start=1
/evidence-not_experimental
/product="anthocyanin 5-aromatic
acyltransferase/benzoyltransferase-like protein"
/protein_id="BAB02518.1"
/db_xref="GI:11994477"
/translation="MTLHVETARVTPTDYSVNSANLKLPLTFDPLWLLFPQVKKR
VFFVELTSTDRDHFHSIILPKLDSLSLILRNYPGLTGWPNPKPSIIIVSENGV
VLVTIAESDADFHLGSGYRGORPLSELHALVPLKPVSDSATFASIQITLFPNGFSIG
VAHHAVLDGKTSFTFKAWAQICKQELQSPENLTPSYDRSLIKKPYTIDERNIELY
RSUKEDVNTKSLTSLFSLSGDGVVLATLVLSRADIERUQKVNKVPSSLHLSTFVI
AYAWFTCFYKRGNGKDRSVLLVFGDFDRDLPKLPGTGYFCNMPVGCYNRKAEE
FMEKGVTAETIISDLVGLSKRIETIADTFVEGFSQSWSTQFTGIAGSTRLGVI
EADFGWRPVKVDIVSIDQGEATAMARRDESQVEIGMCKLTKTMDSVVYFFNGLH
S"
complement(18639..19994)
/note="gene_id:MOD1.4"
/codon_start=1
/evidence-not_experimental
/product="anthocyanin 5-aromatic
acyltransferase/benzoyltransferase-like protein"
/protein_id="BAB02519.1"
/db_xref="GI:11994478"
/translation="MALNVIKISRVSLVNSVEPLVPLTFFDOLLMLKLNPIERVTFY
KLTESSRDSFSLKLEQSLVLSHFLPSGLHKWPNQDPKPHIVIPKPTVSIT
WVSEADFSYISSKELRLETPLPELVQSDSASLSLQITLFPNGFSIGTGVH
HVMYDGKTSKFKHWAHICKHGTPODPLPTVLDRTVINVPAGLEOKIFOLSSYIS
EKDYATPLLPKAKETDNDVVRTLELTVEVDIEKLERAKNESTRSDLLHSTFVYSI
AYVLTKVWCKCGDANRPVRFMAADFRNRDPPVLTIFYGNCVLPIDFNGYKATITFL
GKGYGVNVEILSDSVGLGRNIESIWEYVEDGTKNMKLDTONVTGTSNQRGIYGS
DFGWRPVKVDIVSLYKNNFSSMSARDEIGGLEIGISLKKCMNVFLSLTSDFDIY
K"
complement(27960..28438,28676..28739))
/note="gene_id:MOD1.5
unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB02520.1"
/db_xref="GI:11994479"
/translation="MRLAGACRSPSEESRRRRRARAGWSVDRPIRFIYAADFRNRD
PPVLTIFYGNCVLPIDFNGYKATITFLGEGFVNGVEILSDSVGLGRSLESVWEVIE
EGTKNMKVGTGVLTGSGNQFIYGADFGWRPVNTDMSLYKNNFSSMSARDEIGV
SLKRCRCEMIVFLSLFNGFDN"

```

CDS

```

29217..31674
/note="gene_id:MOD1.6"
/codon_start=1
/pseudo

```

```

/evidence-not_experimental
/product="Mutator-like transposase"
join(32547..32726,32904..33122)
/note="gene_id:MOD1.7"

```

CDS

```

/codon_start=1
/evidence-not_experimental
/product="Mutator-like transposase-like protein"
/protein_id="BAB02521.1"
/db_xref="GI:11994480"

```

```

/translation="MDRLCERDPYDDMKVAKRAIEOMVAMMEGIPKFCPCGGSTV
DTRKDEKRYQCEKEDFTDCHKIRKLWDKAMEEVSLSRESVDYNRKVKLSHEYLI
EEMQKELAHRAEIVNVSKVLFNPNAPKK"
complement(join(33460..33573,33657..34016,34094..34255,
34330..35070,35486..35721,35799..36093,36146..36256))
/note="gb|AAC23763.1
gene_id:MOD1.8

```

CDS

```

strong similarity to unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB02522.1"
/db_xref="GI:11994481"
/translation="MPPKTRGGGKRNTEASAPAKTEKVKAPAEKVKEGAPSKRAK

```

```

VOAPAKKAKVOAPAKKDEKVPAGEOSPTOTATAMATNAAPTAAPTTAPTAPT
ESPLDDSTFYDALKHITPAEIOENMOTDEVEDEETNKELACANVEAEARQDDGL
TVIEEERSESEDEDVNVKESVEDEGEDERDDEIVKSGEERTIDEDIANVDMEE
AMAMQPLMGYFPVSEYPPKMKLATRCYISEVLKTFADLEHLTHKVNFMHPSFKH
IYHLPSTYHKLMMMLFRTKGIIRREDVTKLMEMEPARSKDRLRMAALYFLTSIIVMP
PASYSAGNNMFANRYFKTGVIREDVTKLMEMEPARSKDRLRMAALYFLTSIIVMP
TKTGERASPIDDFCVRSASDLTFCKTFFPSGRYSEYMLKSIHKLDFHNGVNPNTQSP
WVPVGCYPLFLFAFAEAPSLRERFIEEKGAGACGCMCKVKNKRLKMGFFLEQIN
HVLGTTEVIESIIRAEAPSLRERFIEEKGAGACGCMCKVKNKRLKMGFFLEQIN
YEDVQARMEAPNEEVEPTAVGPDPTLVDMMEKLHSINDKLEALLVLMIEBKQAT
FEAFMDQMAKMSQNPDEESTIKENAAAPVVPKRVTRTSKASNVV"
complement(36608..37000)
/note="gene_id:MOD1.9
unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB02523.1"
/db_xref="GI:11994482"

```

CDS

```

/translation="MGGTKASFKYKAHICKHGNIPQDFDLPMLDRTVINVPAGLE
PKLFQLLPYLSKSKVARNMLPPAKENINVRVTLESEANIKKLKEQAKNESTWSD
LLISTFVVTYKSGSLFCKEEAKKSL"
complement(join(40204..40603,40656..40872))
/note="gb|AAC97247.1
gene_id:MOD1.10
similar to unknown protein"
/codon_start=1
/pseudo

```

CDS

```

/evidence-not_experimental
complement(join(42135..42706,42772..43162))
/note="gene_id:MOD1.11
unknown protein"
/codon_start=1

```

```

/evidence-not_experimental
/protein_id="BAB02524.1"
/db_xref="GI:11994483"

```

CDS

```

/translation="MALNMKISRSHATNSVTPFLPTFFDOLLMLKHHPIERVTFY
KLSESSHDSFFSTILPKLQSLALTNFLPLSGHKKWPDQPKPHIVIPKPTVSS
LWLSPLTSDSDSPAKSLVSRSYVLLSLRSLGTTTHHVIMDKSTNFNFKHSAHICK
LVRTIPLQGVLPVTHSDRTVINVSGLFEKILQLLSLGSVDNDRTILKSPAKEDIDD
LVRVTIQLSQENIKKIKERTKNESTRDLSLTFVVKYAYVLTCVVKARGNVDPRPIR
FMYAADFNRNRDPPVLTIFYGNCVLPIDIFHGRRWCEWGRGSK"
44134..44573
/note="gene_id:MOD1.12
unknown protein"
/codon_start=1
/pseudo

```

CDS

```

/evidence-not_experimental

```



```
CDS      complement(46109..47990)
          /note="contains similarity to reverse transcriptase
          gene_id:MOD1.13"
          /codon_start=1
          /pseudo
          /evidence=not_experimental
CDS      complement(join(48707..49066,49148..49372,49477..49764,
          49850..49982,50772..51282))
          /note="gb|AAAF06088.1
          gene_id:MOD1.14
          similar to unknown protein"
          /codon_start=1
          /pseudo
          /evidence=not_experimental
          complement(52336..56745)

Query Match      5.1%; Score 77.8; DB 12; Length 85690;
Best Local Similarity 48.5%; Pred. No. 7.6e-06;
Matches 276; Conservative 0; Mismatches 287; Indels 6; Gaps 2;

QY  910 GATTTAAGCGCGTTGGATCCGCGGCTCCGGGAACACTTCTCGAAACTGCTCTATCG 969
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  16051 GATTTTCAGAGACCGGTTAGATCCGAAGCTCCCGGGGACTTACTTCGGAACCTGATGATT 16110

QY  970 TTTTCGATGCCGAAGATCCTCGCGCGGATTTGGTCGGAGATGAAGGGGTGTTTCGGGCA 1029
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  16111 CCGGTGGCTGTTATACCGTAGAGCGCGGAGTTTATGAGAGAGAGGATTTGTGACG 16170

QY  1030 GCTGAGCGCATCCGCGGGAATAGAGAAGCAGCAGGCGACAGAAGATCTTACGAAACT 1089
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  16171 GCGGCTGAGATTATTAGCGATTGTTGTAAGGCTTCAGTTCGAAGAAGATAGAGACCA 16230

QY  1090 GTGAGAACTGGCCGCTGAGATTCGCG---AAGCCCTGCAAAACTGTTATTTCTCGGTG 1146
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  16231 GCGGACACTTTCGTGAAGGATTTAGTTTTCAGAGTTGGAGTACACAGTTTGGGACGATA 16290

QY  1147 GCGGGATCAGCAGCGCTGATCTTTACGGCGCGGATTTTGGATGGGTAAAGCGGTGAAG 1206
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  16291 GCGGGTCGACCCGTTGGGAGTATACGAGCGGATTTCCGGTGGGAAGACCCGTTAA 16350

QY  1207 CAAGATACCTGTCGATGATGAGAGAGTTTACGATGTCGTTGTGTAACCCGAGGGAT 1266
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  16351 GTTGATATTGTCCTCCATGATACCAAGAGAGA---AGCGATCGCAATGGCTGAGAGACG 16407

QY  1267 GCTGCGGAGGATTTGGAGTTGATGTCCTTCCCAAGGAGCAATTCGAAGCTTTTGAT 1326
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  16408 GACTAGGTGGCGTTGAGATTGGAATGTGTTGAAAAGACTGAAATGGATTCGCGTCGTG 16467

QY  1327 GATTATTTTCGGAGGGAATAAAGGGTTGATTAATCATTTTAATCATGATATATGAAGTTG 1386
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  16468 TCTTTTTCACAATGGTTTACATAGCTAAACTCTGTTTGTGTAAGATTTTAATAATAAG 16527

QY  1387 GATGAATCCTCTGTTTCATCTCTATTGTTTAAACAATAATTTTTTTTCCATTTGAAC 1446
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  16528 GTAAATGCTGGTTAGTATTCTTGTGTTTTTTTCTCACAAGAAGATATGGGCATAATTCT 16587

QY  1447 TTCAGTCAATAAATAAATAAATAAATAA 1475
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  16588 ACTAATAAATAAGTTTAAATAAAGATATA 16616
```

```
RESULT 12
AC069470
LOCUS      AC069470 136047 bp DNA HTG 19-JAN-2001
DEFINITION Arabidopsis thaliana chromosome I clone IGF-F7F7, *** SEQUENCING IN
            PROGRESS ***, 4 unordered pieces.
ACCESSION AC069470
VERSION   AC069470.10 GI:12320593
KEYWORDS  HTG: HTGS_PHASE1.
SOURCE    thale cress.
           Arabidopsis thaliana
           Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
           Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
```

```
Brassicales: Brassicaceae; Arabidopsis.
1 (bases 1 to 136047)
Lin.X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Ronning,C.M.,
Koo,H., Fujii,C.Y., Utterback,T., Barnstead,M.E., Bowman,C.L.,
White,O., Nierman,W.C. and Fraser,C.M.
Arabidopsis thaliana 'IGF' BAC 'F7F7' genomic sequence near marker
'mi72'
Unpublished
2 (bases 1 to 136047)
Town,C.D. and Kaul,S.
Direct Submission
Submitted (01-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org
On Jan 19, 2001 this sequence version replaced gi:12280790.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 100163: contig of 100163 bp in length
* 100164 100213: gap of unknown length
* 100214 112623: contig of 12410 bp in length
* 112624 112673: gap of unknown length
* 112674 116636: contig of 3963 bp in length
* 116637 116886: gap of unknown length
* 116887 136047: contig of 19361 bp in length.
FEATURES             Location/Qualifiers
     source            1..136047
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="I"
                     /clone="IGF-F7F7"
BASE COUNT  46028 a 22872 c 22213 g 44783 t 151 others
ORIGIN

Query Match      5.1%; Score 77.8; DB 74; Length 136047;
Best Local Similarity 48.5%; Pred. No. 8e-06;
Matches 276; Conservative 0; Mismatches 287; Indels 6; Gaps 2;

QY  910 GATTTAAGCGCGGTTGGATCCGCGGCTCCGGGAACACTTCTCGAAACTGCTCTATCG 969
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  90952 GATTTTCAGAGACCGGTTAGATCCGAAGCTCCCGGGGACTTACTTCGGAACCTGATGATT 91011

QY  970 TTTTCGATGCCGAAGATCCTCGCGCGGATTTGGTCGGAGATGAAGGGGTGTTTCGGGCA 1029
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  91012 CCGGTGGCTGTTTATAACCCTAAGCGCGGAGTTTATGGAAGAGAGGATTTGTGACG 91071

QY  1030 GCTGAGCGCATCCGCGGGAATAGAGAAGCAGCAGGCGACAGAAGATCTAGAAACT 1089
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  91072 GCGGCTGAGATTATTACGCAATTTGGTGAAGGGTTGAGTTCAAGAAGATAGAGACCA 91131

QY  1090 GTGAGAACTGGCCGCTGAGATTCGCG---AAGCCCTGCAAAACTGTTATTTCTCGGTG 1146
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  91132 GCGGACACTTTCGTGAAGGATTTAGTTTTCAGAGTTGGAGTACACAGTTTGGACGATA 91191

QY  1147 GCGGGATCAGCAGCGCTTGA TCTTTACGCGCGGATTTTGGATGGGTAAAGCGGTGAAG 1206
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  91192 GCGGGTCGACCCGTTGGGAGTATACGAGGCGGATTTCCGGTGGGAAGACCCGTTAA 91251

QY  1207 CAAGATACCTGTCGATGATGAGAGAGATTTACGATGTCGTTGTGTAACCCGAGGGAT 1266
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  91252 GTTGATATTGCTCTCCATTTGACCAAGGAGA---AGCGATCGCAATGGCTGAGAGACG 91308

QY  1267 GCTGCGGAGGATTTGGAGGTCGATTGCTTCCAAAGGAGGAATTCGAAGCTTTTGAT 1326
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db  91309 GAGTCAGGTGGCGTTGAGATTGGAATGTGTTGAAAAGACTGAATGGATTCGCGTCGTG 91368

QY  1327 GATTATTTTCGGAGGGAATAAAGGGTTGATTAATCATTTTAATCATGATATATGAAGTTG 1386
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 91369 TCTTTTTCACAAATGGTTTACATAGCTAAACTCTTGTGTTGTAAGATTTTAATAATAAG 91428
Qy 1387 GATGAATCCTCTGTTTCATCTCTATGTTTAAACAATAAATTTTTTCCATTTGAACCTTTT 1446
Db 91429 GTAATGCTTGGTTAGTTATCTGTTTCTCACAAGAAGTATGGGCATATTCCTCC 91488
Qy 1447 TTGAGTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1475
Db 91489 ACTAATAATAAGTTTAAATAAAGATATA 91517

RESULT 13
AX025514
LOCUS AX025514 1648 bp DNA PAT 16-SEP-2000
DEFINITION Sequence 40 from Patent WO0032789.
ACCESSION AX025514
VERSION AX025514.1 GI:10187184
KEYWORDS Citrus limon.
SOURCE Citrus limon.
ORGANISM Citrus limon.
REFERENCE 1 (bases 1 to 1648)
AUTHORS Aharoni,A., Verhoeven,H.A., Luecker,J., O'Connell,A.P. and Van Tunen,A.J.
TITLE Fruit flavour related genes and use thereof
JOURNAL Patent: WO 0032789-A 40 08-JUN-2000.
AHARONI ASAPH (IL) ; VERHOEVEN HARRIE ADRIANUS (NL) ; LUECKER JOOST (NL) ; CPRO DLO (NL) ; CONNELL ANN PATRICIA O (NL) ; TUNEN ARJEN JOHANNES VAN (NL)
FEATURES
source Location/Qualifiers
1. .1648
/organism="Citrus limon"
/db_xref="taxon:2708"
52. .1527
/note="cDNA"
Lemon acyl transferase"
/codon_start=1
/protein_id="CAC09066.1"
/db_xref="GI:10187185"
/translaton="MAAIENKRVLLKHEVTKVTPFVNPNSKTSFTLIDLTYFFDFWPK
NPVERLFFVDFDLNENSEILPKLHSLFTLLHLYPLAGHIMPLDAAKPAVY
YFPDQGVFAVAENWSECHGPHLSGIRQAVEFHLVPLQSLSDKAEVIAIQ
ITLFPNGGFSIGVSHHAILDGGKSTFLKSWAYLCKLOLCHHPCLSPDLTLLDRT
VIKDPTGDMLOLNKVVGVGSDNSDPQKIRSLKVLPLFUDSESLNKLVRATFELREDIT
KLIRKVNHOLSKSKSKQVRLSTFVLFLAYFVCMKAKAKAKATEAAAGANDEIKN
IIVGFTADYRSRLDPLPLNFGMNCGRHETAKASDFVQENGFAVFAEMLSDMVKI
DADAIEDDKVSEILKELKGAIFSVAGSTQFDVYGSDFGMRPKKVEIVSIDRTQ
AISAERDDGGGVEGVVLEKQOMEVFESVFADGLKNDLV"
BASE COUNT 464 a 344 c 354 g 486 t
ORIGIN

Query Match 4.7%; Score 71.2; DB 9; Length 1648;
Best Local Similarity 56.2%; Pred. No. 0.00012;
Matches 199; Conservative 0; Mismatches 143; Indels 12; Gaps 3;

Qy 1137 TTTCTCGTGGCGGATCGACGAGCTTGATCTTTACGGCGCGATTTTGGATGGGTAA 1196
Db 1302 TTTTCTCTGCTGCTCGACCCCAATTTGATGTTTACGGTTCGGATTTTGGGTGGGGAG 1361
Qy 1197 GCGGTGAAGCAAGATACACTGTCGATTGATGGAGAGATTTTACGATGCTGTTGTGTA 1256
Db 1362 GCCCAAGAGGTGGAGATTGTCAATAGATAG---GACACAGCCATCTCTTTGGCAGA 1418
Qy 1257 ACCGAGGATGCTCGCGGAGGATGGAGTTGATTTGCTTTGCCAAGAGGAGGATGCA 1316
Db 1419 GAGAAGAGATGGAGGCGGCTTGAGGTTGGAGTTCTTTTAGAGAGCAACAATGGA 1478
Qy 1317 AGCTTTTGATGATTATTTTGGGAGGAAT-AAAGGGTTGATTATCATTTTAATCATGTA 1375
Db 1479 GGTTTTGAATCTGATTGCTGATGGAGTGAATAAATGATCTTGTGTTAATTAATGATGA 1538
```

```
Qy 1376 TTATGAAGTTGGATGAATCCTCTGTTTTCATCTCTATGTTTAAACAATAAATTTTTTCC 1435
Db 1539 TCAT-----CTAAATTTCTCAATATATTATTTGGTCATATTCAAAGAAATAAATTAT 1590
Qy 1436 ATTGAACHTTTTGGTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1489
Db 1591 TCGGATTTTGTGACCAACCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1644

RESULT 14
CNS07360
LOCUS CNS07360 1167 bp DNA STS
DEFINITION clone BA0AB017A09 of library BA0AB from strain CLIB 210 of Kluyveromyces lactis, sequence tagged site.
ACCESSION AL427102
VERSION AL427102.1 GI:12210296
KEYWORDS STS.
SOURCE Kluyveromyces lactis.
ORGANISM Kluyveromyces lactis.
REFERENCE 1 (bases 1 to 1167)
AUTHORS Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F., Durrans,P., de-Montigny,J., Dujon,B., Robert,C., Terrier,M., Wincker,P., Saurin,W., Tekala,F., Wincker,P. and Weissbach,J.
TITLE Genomic Exploration of the Hemiascomycetous Yeasts: 11.
JOURNAL FEBS Lett. 487 (1), 66-70 (2000)
PUBMED 11152886
REFERENCE 2 (bases 1 to 1167)
AUTHORS Souciet,J.L., Algle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrans,P., Lepingle,A., Llorente,B., Maupertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissbach,J.
TITLE Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of yeast species for molecular evolution studies(1)
JOURNAL FEBS Lett. 487 (1), 3-12 (2000)
PUBMED 11152876
REFERENCE 3 (bases 1 to 1167)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (08-Sep-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT This STS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
FEATURES
source Location/Qualifiers
1. .1167
/organism="Kluyveromyces lactis"
/strain="CLIB 210"
/variety="lactis"
/db_xref="taxon:28985"
/clone="BA0AB017A09"
/clone_lib="BA0AB"
1. .1167
/note="part of mitochondrial DNA"
/evidence="not_experimental"
BASE COUNT 335 a 117 c 96 g 498 t 121 others
ORIGIN

Query Match 4.6%; Score 70.2; DB 53; Length 1167;
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:11:58 ; Search time 445.49 Seconds
(without alignments)
2145.203 Million cell updates/sec

Title: US-08-894-356C-6
Perfect score: 1522
Sequence: 1 NTGACCACCCTCGAATC.....TTTTTTTTTTTTTTTTTTTT 1522

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Maximum Match 0%
Listing first 45 summaries

Database : N_Geneseq_0601.*
1: /SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1505	98.9	1518	17	AAT37313 Aromatic acyl tran
2	765.4	50.3	1479	17	AAT37311 Aromatic acyl tran
3	196.2	12.9	1703	17	AAT37308 Aromatic acyl tran
4	174.2	11.4	1508	17	AAT37312 Aromatic acyl tran
5	153.8	10.1	1622	17	AAT37309 Aromatic acyl tran
6	87	5.7	936	22	AAF58252 Oligonucleotide D1
7	87	5.7	936	22	AAF58254 Oligonucleotide D1
8	87	5.7	936	22	AAF58257 Oligonucleotide D1
9	87	5.7	936	22	AAF58259 Oligonucleotide D1
10	87	5.7	936	22	AAF58262 Oligonucleotide D2
11	87	5.7	938	22	AAF58255 Oligonucleotide D1

12	80.2	5.3	936	22	AAF58252 Oligonucleotide D1
13	80.2	5.3	936	22	AAF58254 Oligonucleotide D1
14	80.2	5.3	936	22	AAF58257 Oligonucleotide D1
15	80.2	5.3	936	22	AAF58259 Oligonucleotide D2
16	80.2	5.3	936	22	AAF58262 Oligonucleotide D2
17	80.2	5.3	938	22	AAF58255 Oligonucleotide D1
18	72.4	4.8	244	22	AAF58238 Oligonucleotide D1
19	72.4	4.8	244	22	AAF58238 Oligonucleotide D1
20	71.2	4.7	1648	21	AAC64785 Lemon acyl transe
21	65	4.3	1240	21	AAC53227 Arabidopsis thalia
22	62.4	4.1	49999	20	AAC23902 Human LOBO homolo
23	61.6	4.0	1379	21	AAC46852 Arabidopsis thalia
24	55.4	3.6	49999	20	AAC23902 Human LOBO homolo
25	55	3.6	2418	13	AAQ27886 P.falciplarum GBp13
26	54.6	3.6	3030	21	AAC61886 cDNA encoding a hu
27	54.2	3.6	1520	21	AAF15829 Human prostate can
28	54	3.5	2662	10	AAQ07073 Rhoptry membra
29	52.6	3.5	1875	18	AAT48669 Human EDG-2 recept
30	51.8	3.4	1213	21	AAC51043 Arabidopsis thalia
31	51.8	3.4	1217	21	AAC42550 Arabidopsis thalia
32	51.4	3.4	3138	12	AAQ11712 Shuttle vector PMU
33	51	3.4	8365	20	AAZ20056 Plasmodium falcipa
34	50.8	3.3	1052	10	AAQ90224 Malaria-specific D
35	50	3.3	1518	17	AAT37313 Aromatic acyl tran
36	50	3.3	5173	18	AAT89783 Human Immunodefici
37	49.8	3.3	1132	20	AAQ80668 Clone am996_12 enc
38	49.8	3.3	2229	21	AAQ23425 cDNA encoding huma
39	49.6	3.3	3138	12	AAQ11712 Shuttle vector PMU
40	49	3.2	2564	22	AAQ00034 Human ATLAS-3-enco
41	49	3.2	4590	7	AAQ60472 Sequence encoding
42	48.8	3.2	2418	13	AAQ27886 P.falciplarum GBp13
43	48.6	3.2	2535	21	AAF18128 Lung cancer associ
44	48.4	3.2	1493	11	AAQ04492 Sequence encoding
45	48.4	3.2	1493	20	AAQ33994 Human osteopontin

ALIGNMENTS

RESULT 1
AAT37313
ID AAT37313 standard; cDNA to mRNA; 1518 BP.
XX
AC AAT37313;
XX
DT 06-FEB-1997 (first entry)
XX
DE Aromatic acyl transferase coding sequence.
XX
KW Aromatic acyl transferase; transformation; anthocyanin pigment;
KW plants; acylation; colour; tone; colouration; colour change;
KW Gentiana triflora; Petunia hybrida; Perilla oclmoides;
KW Scenecio cruentus; Lavandula angustifolia; ds.
XX
OS Lavandula angustifolia (Clone pLAT21).
XX
FH Key Location/Qualifiers
FT CDS l..1355
FT /*tag- a
FT /product= Aromatic acyl transferase.
XX
XX WO9625500-A1.
XX
XX 22-AUG-1996.
XX
XX 16-FEB-1996; 96WO-JP00348.
XX
XX 30-JAN-1996; 96JP-0046534.
XX
XX 17-FEB-1995; 95JP-0067159.
XX
XX 29-JUN-1995; 95JP-0196915.
XX
XX (SUNR) SUNTORY LTD.

Oy	1207	CAAGAGATACTGTCGATTGATCGAGAGAAGTTTACGATGTGGTTGTAAACCCAGGGAT	1266
Dg	1263	tctgacattaccctcgttgattatgcagaatt--gatttatgtgattccagtcagggat	1319
Oy	1267	GCTGCCGGAGGATTGCAGGTTCGATTGCTTCTTGCCAAAGGAGGAATTCCAAGCTTTTGAT	1326
Dg	1320	tctgaaaaaggctggagagtggatcatcgctgaagattcatatggatgcatttgcga	1379
Oy	1327	GATTATTTCGCGAGGG	1343
Dg	1380	aaatctttgaagaagg	1396
RESULT	4		
ID	AAT37312		
XX	AAT37312 standard; cDNA to mRNA; 1508 BP.		
AC	AAT37312;		
XX			
DT	06-FEB-1997 (first entry)		
XX	Aromatic acyl transferase coding sequence.		
XX			
KW	Aromatic acyl transferase; transformation; anthocyanin pigment;		
KW	plants; acylation; colour; tone; colouration; colour change;		
KW	Gentiana triflora; Petunia hybrida; Perilla ocimoides;		
KW	Scenecio cruentus; Lavandula angustifolia; ds.		
XX			
OS	Senecio cruentus (Clone pCAT8).		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	3..1367	
FT		/*tag= a	
FT		/product= Aromatic acyl transferase.	
XX	WO9625500-A1.		
XX			
PD	22-AUG-1996.		
XX			
PF	16-FEB-1996; 96WO-JP00348.		
XX			
XX	30-JAN-1996; 96JP-0046534.		
PR	17-FEB-1995; 95JP-0067159.		
PR	29-JUN-1995; 95JP-0196915.		
XX			
PA	(SUNR) SUNTORY LTD.		
XX			
PI	Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;		
PI	Nakao M, Tanaka Y, Yonekura K;		
XX			
WI	WPI; 1996-393401/39.		
DR	P-PSDB; AAW04726.		
XX			
PT	DNA coding for aromatic acyl transferase - for transforming plants		
PT	which produce anthocyanin pigments and thus altering colour tone,		
PT	e.g. of flowers		
XX			
PS	Claim 4; Page 69-72; 94pp; Japanese.		
XX			
CC	Vectors containing DNA fragments encoding proteins of plant origin		
CC	with aromatic acyl transferase activity may be used to transform		
CC	plants which produce anthocyanin pigments. The aromatic acyl		
CC	transferase acylates the pigments in the flower resulting in colour		
CC	tone changes and allowing new colourations to be produced. Six		
CC	specific DNA sequences encoding aromatic acyl transferase from		
CC	different plants are described in AAT37308-T37313.		
XX			
SQ	Sequence 1508 BP; 442 A; 294 C; 295 G; 477 T; 0 other;		

Query Match 11.4%; Score 174.2; DB 17; Length 1508;
Best Local Similarity 48.3%; Pred. No. 1.1e-24;


```

Db 710 gaagaaactgctcaaatctcaggccacactactgttcttaaatccagcaatttctaaa 769
Qy 709 GATCGGATTGCGAGCTACCTACCTTTTCACCCCAATCTGAATTAAGAAATGGAAGGTTG 768
Db 770 gatgaagtctgagccactctcctacacccctattgatcatgatgaagctcagaataatc 829
Qy 769 ATTACAGAGAAAGCCCAAAATGTA-----GTTAATCTCTCTTCCCTTCGTC 813
Db 830 atttcgtcaaaaaatcgcaacttaaccggtagtagtaattataatctgtcaactttcacg 889
Qy 814 GCGATCGCAGCTTATCTGACCGCCGATCGCCAAAT-----CGGTGCGAGATTAC 864
Db 890 gfgacatctgcactgatctcggacatgcttgcgaataatttagacacccgtcgaagag 949
Qy 865 AAAGACGTGGATGAGACAAACCGCTTCTTTTAAATTCGATCGATTAAGCGCGGT 924
Db 950 aaggtggaagagataaaacatcagcaaatctgtctttcatcaactgccgaacagt 1009
Qy 925 TTGGATCCCGCGCTCCGGGGAACACTTCCGAAACTGTCTATCGTTTCGATGCGGAAG 984
Db 1010 ttgtcgcgcatacctcaaaattactttggaattgcatagtccttgatggtgga 1069
Qy 985 ATCTCGCGCGGATTTGGTCGAGATGAAGGGGTGTTTCGCGCAGCTGAGCGATCGCG 1044
Db 1070 tgcactcatgagcaactttaggaaatgaagggttgcgtgtagtgcaaccgccatcgga 1129
Qy 1045 GCGGAATAGACGAAGGAGCGACGACACAGAA---GATTCTAGAAACTGTGAGAACTGG 1101
Db 1130 gatctatccataaagggtttacatgactacgaaagaaatctgagaggagattggtatcg 1189
Qy 1102 CGCTCTGAGATTCCGCAAGCCTTGCAAAACT-----GTTATTCTCGGTGCGCGGA 1152
Db 1190 cgcgccgatacaactctcggcaccagaaggtcgacgctcatattatgcttbgatccgca 1249
Qy 1153 TCGAGACGCTTGATCTTTTACCGCGGATTTTGGATGGGGTGAAGCGGTGAAGCAAGAG 1212
Db 1250 caacgcaatgcatgattttgacgatttctggtggtggaagcttgaaagcatgaa 1309
Qy 1213 ATACTGTCGATGATGAGAGAGAGTTTACGATGTCGTTGTGTAACCGAGGATGCTGCC 1272
Db 1310 tctgtttcaactcaactccttcggcaccactaatatttgatctctcgtccagaagatttaa 1369
Qy 1273 GGAGGATTGAGGTTGGATTGCTTTTGCCAAAGGAGGAATTCGAAGCTTTTGATGATTAT 1332
Db 1370 gtagcacttgacttggcattcttcttgcctaagaataggatggacgcatcttgcaccatt 1429
Qy 1333 TTTGCGGAGGAATAAAGGGT 1353
Db 1430 tttagcaatttcatcaatagt 1450

```

RESULT 6

```

AAF58252/c
ID AAF58252 standard; DNA; 936 BP.
XX
AC AAF58252;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1835.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000WO-US20476.
XX

```

```

PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX Umek RM;
XX WPI; 2001-159728/16.
XX
PT Nucleic acids containiny electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

```

```

Query Match 5.7%; Score 87; DB 22; Length 936;
Best Local Similarity 6.2%; Pred. No. 4.3e-08;
Matches 33; Conservative 295; Mismatches 205; Indels 0; Gaps 0;
Qy 990 GCGGCGGGATTTGGTCGAGATGAAGGGGTGTTTCGCGCAGCTGAGCGCATCGGCGGA 1049
Db 844 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 785
Qy 1050 AATAGAGAAGAGGACGACGACAGAAAGATTCTAGAAACTGTGGAACACTGGCCGTCTGA 1109
Db 784 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 725
Qy 1110 GATTCGCGAAGCTTGCAAACTGTATTCTCGGTGCGCGGATCGACGAGCTTCATCT 1169
Db 724 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 665
Qy 1170 TTACGCGCGGATTTTGGATGGGGTGAAGCGGTGAAGCAAGAGACTCTCGATGATGG 1229
Db 664 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 605
Qy 1230 AGAGAAGTTTACGATGCTCTGTGTAAACCGAGGAGTCTCGCGGAGGATTCGAGTTGG 1289
Db 604 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 545
Qy 1290 ATTGCTCTTGCACAAAGGAGGAATTGCAAGCTTTTGATGATTATTTTTCGCGGGAATAA 1349
Db 544 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 485
Qy 1350 GGGTTGATTATTAATCATATGATGATGATGATGATGATGATGATGATGATGATGATG 1409
Db 484 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 425
Qy 1410 TATGTTTAAACAATAATTTTTCATTCGAACATTTTTCAGTCAATAAAAAA 1469
Db 424 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 365
Qy 1470 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1522
Db 364 WWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWWW 312

```

RESULT 7

```

AAF58254/c
ID AAF58254 standard; DNA; 936 BP.
XX
AC AAF58254;
XX
DT 24-APR-2001 (first entry)
XX

```


KW Electron-transfer group; ETM; mismatch; genotyping;
XX gene expression; ss.
OS Synthetic.
XX WO200107665-A2.
PN 01-FEB-2001.
XX 26-JUL-2000; 2000WO-US20476.
XX 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX (CLIN-) CLINICAL MICRO SENSORS INC.
PA Umek RM;
XX WPI; 2001-159728/16.
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX Example 6; Page 127; 159pp; English.
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

DE Oligonucleotide D1875.
XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX Synthetic.
XX OS
XX PN WO200107665-A2.
XX PD 01-FEB-2001.
XX 26-JUL-2000; 2000WO-US20476.
XX 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX (CLIN-) CLINICAL MICRO SENSORS INC.
PA Umek RM;
XX WPI; 2001-159728/16.
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX Example 6; Page 127; 159pp; English.
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 5.3%; Score 80.2; DB 22; Length 936;
Best Local Similarity 1.3%; Pred. No. 8.6e-07;
Matches 4; Conservative 219; Mismatches 92; Indels 0; Gaps 0;
QY 1208 AAGAGATACGTGCGATTGGAGAGAGTTTACGATGTCGTTGTGTAACCGAGGATG 1267
DB 249 AAGAGATACGTGCGATTGGAGAGAGTTTACGATGTCGTTGTGTAACCGAGGATG 308
QY 1268 CTCCCGGAGGATTGGAGGTTGGATTGCTTTGCCAAGGAGGAATGCCAAGCTTTTGATG 1327
DB 309 CTCCCGGAGGATTGGAGGTTGGATTGCTTTGCCAAGGAGGAATGCCAAGCTTTTGATG 368
QY 1328 ATTATTTTGGCGAGGAATAAGGGTTGATTATCATGTTATTAATCATGTTGAGTTGG 1387
DB 369 ATTATTTTGGCGAGGAATAAGGGTTGATTATCATGTTATTAATCATGTTGAGTTGG 428
QY 1388 ATGAATCCCTGCTGTCATCTCTATGTTTAAACAATAATTTTTCATGTAACCTTTT 1447
DB 429 ATGAATCCCTGCTGTCATCTCTATGTTTAAACAATAATTTTTCATGTAACCTTTT 488
QY 1448 TGAGTCAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1507
DB 489 TGAGTCAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 548
QY 1508 TTTTNTTTTTTTTTT 1522
DB 549 TTTTNTTTTTTTTTT 563

Query Match 5.3%; Score 80.2; DB 22; Length 936;
Best Local Similarity 1.3%; Pred. No. 8.6e-07;
Matches 4; Conservative 219; Mismatches 92; Indels 0; Gaps 0;
QY 1208 AAGAGATACGTGCGATTGGAGAGAGTTTACGATGTCGTTGTGTAACCGAGGATG 1267
DB 249 AAGAGATACGTGCGATTGGAGAGAGTTTACGATGTCGTTGTGTAACCGAGGATG 308
QY 1268 CTCCCGGAGGATTGGAGGTTGGATTGCTTTGCCAAGGAGGAATGCCAAGCTTTTGATG 1327
DB 309 CTCCCGGAGGATTGGAGGTTGGATTGCTTTGCCAAGGAGGAATGCCAAGCTTTTGATG 368
QY 1328 ATTATTTTGGCGAGGAATAAGGGTTGATTATCATGTTATTAATCATGTTGAGTTGG 1387
DB 369 ATTATTTTGGCGAGGAATAAGGGTTGATTATCATGTTATTAATCATGTTGAGTTGG 428
QY 1388 ATGAATCCCTGCTGTCATCTCTATGTTTAAACAATAATTTTTCATGTAACCTTTT 1447
DB 429 ATGAATCCCTGCTGTCATCTCTATGTTTAAACAATAATTTTTCATGTAACCTTTT 488
QY 1448 TGAGTCAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1507
DB 489 TGAGTCAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 548
QY 1508 TTTTNTTTTTTTTTT 1522
DB 549 TTTTNTTTTTTTTTT 563

RESULT 13
ID AAF58254
AC AAF58254 standard; DNA; 936 BP.
XX AAF58254;
XX 24-APR-2001 (first entry)
XX

RESULT 14
ID AAF58257
AC AAF58257 standard; DNA; 936 BP.
XX AAF58257;
XX

DT 24-APR-2001 (first entry)
 DE Oligonucleotide D1954.
 DE Electron-transfer group; ETM; mismatch; genotyping;
 KW gene expression; ss.
 XX Synthetic.
 XX WO200107665-A2.
 XX 01-FEB-2001.
 XX 26-JUL-2000; 2000WO-US20476.
 XX 26-JUL-1999; 99US-0145695.
 PR 17-MAR-2000; 2000US-0190259.
 XX (CLIN-) CLINICAL MICRO SENSORS INC.
 PA Umek RM;
 PI WPI; 2001-159728/16.
 DR Nucleic acids containing electron-transfer group, useful as labels in
 PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
 PT a single surface -
 PS Example 6; Page 127; 159pp; English.
 XX The present invention relates to a composition comprising two nucleic
 CC acids each containing an electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 CC detection of nucleic acids, especially of substitutions (mismatches)
 CC and single-nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.
 XX Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

Query Match 5.3%; Score 80.2; DB 22; Length 936;
 Best Local Similarity 1.3%; Pred. No. 8.6e-07;
 Matches 4; Conservative 219; Mismatches 92; Indels 0; Gaps 0;
 Qy 1208 AAGAGATACTGCGATTGAGAGAGAGTTTACGATGCTGTGTAACCGAGGGATG 1267
 Db 249 www. 308
 Qy 1268 CTGCCGAGGATTGGAGTTGCTTTGCCAAAGAGGAATTGCAAGCTTTTGATG 1327
 Db 309 www. 368
 Qy 1328 ATTATTTTCGAGGAATAAAGGGTTGATTAATCAATTAATCAATTAATGAAGTTGG 1387
 Db 369 www. 428
 Qy 1388 ATGAATCCTCTGTTTCATCTATTTGTTAAACAATAATTTTCCATTCGAATTTT 1447
 Db 429 www. 488
 Qy 1448 TGAGTCAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1507
 Db 489 www. 548
 Qy 1508 TTTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1522
 Db 549 www. 563

RESULT 15
 AAF58259
 ID AAF58259 standard; DNA; 936 BP.
 XX

AC AAF58259;
 XX 24-APR-2001 (first entry)
 XX Oligonucleotide D2004.
 XX Electron-transfer group; ETM; mismatch; genotyping;
 KW gene expression; ss.
 XX Synthetic.
 XX WO200107665-A2.
 XX 01-FEB-2001.
 XX 26-JUL-2000; 2000WO-US20476.
 XX 26-JUL-1999; 99US-0145695.
 PR 17-MAR-2000; 2000US-0190259.
 XX (CLIN-) CLINICAL MICRO SENSORS INC.
 PA Umek RM;
 PI WPI; 2001-159728/16.
 DR Nucleic acids containing electron-transfer group, useful as labels in
 PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
 PT a single surface -
 PS Example 6; Page 128; 159pp; English.
 XX The present invention relates to a composition comprising two nucleic
 CC acids each containing an electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 CC detection of nucleic acids, especially of substitutions (mismatches)
 CC and single-nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.
 XX Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

Query Match 5.3%; Score 80.2; DB 22; Length 936;
 Best Local Similarity 1.3%; Pred. No. 8.6e-07;
 Matches 4; Conservative 219; Mismatches 92; Indels 0; Gaps 0;
 Qy 1208 AAGAGATACTGCGATTGAGAGAGAGTTTACGATGCTGTGTAACCGAGGGATG 1267
 Db 249 www. 308
 Qy 1268 CTGCCGAGGATTGGAGTTGCTTTGCCAAAGAGGAATTGCAAGCTTTTGATG 1327
 Db 309 www. 368
 Qy 1328 ATTATTTTCGAGGAATAAAGGGTTGATTAATCAATTAATCAATTAATGAAGTTGG 1387
 Db 369 www. 428
 Qy 1388 ATGAATCCTCTGTTTCATCTATTTGTTAAACAATAATTTTCCATTCGAATTTT 1447
 Db 429 www. 488
 Qy 1448 TGAGTCAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1507
 Db 489 www. 548
 Qy 1508 TTTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1522
 Db 549 www. 563

Search completed: November 5, 2001, 18:12:08

Job time: 15463 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:04:56 ; Search time 168.74 seconds
(without alignments)
1707.547 Million cell updates/sec

Title: US-08-894-356C-6

Perfect score: 1522

Sequence: 1 NTGACCACCTCTCGAATC.....TTTTTTTTTTTTTTTTTTTT 1522

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/2/lna/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/lna/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/lna/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/lna/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/lna/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/lna/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	3.7	7218	1	US-08-232-463-14
2	56.8	3.7	7218	1	US-08-232-463-14
3	52.6	3.5	1875	5	PCT-US96-10618-1
4	51.4	3.4	3138	1	US-07-867-106-4
5	50	3.3	5173	1	US-08-242-677-1
6	49.8	3.3	130	6	5198345-15
7	49.8	3.3	130	6	5198345-15
8	49.6	3.3	3138	1	US-07-867-106-4
9	49.6	3.3	5852	4	US-07-867-106-2
10	49.2	3.2	1578	4	US-09-416-050A-1
11	49.2	3.2	1578	4	US-09-664-800-1
12	49.2	3.2	1578	4	US-09-665-309-1
13	49.2	3.2	1578	4	US-09-661-569-1
14	48.4	3.2	1493	6	5340934-5
15	47.6	3.1	198	5	PCT-US92-10087-16
16	47.6	3.1	198	5	PCT-US92-10087-16
17	47.2	3.1	5852	1	US-07-867-106-2
18	46.8	3.1	860	1	US-07-847-010-18
19	46.8	3.1	4818	3	US-08-817-928-27
20	46.6	3.1	1582	3	US-08-545-196B-10
21	46.6	3.1	1582	3	US-08-545-196B-12
22	46.6	3.1	22481	5	PCT-US95-07201-43
23	46.4	3.0	2836	3	US-08-747-221B-24
24	46.4	3.0	2836	3	US-08-747-221B-26
25	45.6	3.0	3095	6	5231168-1
26	45.4	3.0	2209	1	US-08-514-014-1
27	45.4	3.0	2209	2	US-08-833-823-1

```

28 45.4 3.0 2280 3 US-08-813-150-1
29 44.6 2.9 860 1 US-07-847-010-18
30 44.6 2.9 2422 1 US-07-867-106-5
31 43.8 2.9 222 4 US-08-481-190-15
32 43.8 2.9 222 5 PCT-US93-00889-15
c 33 43.6 2.9 83 1 US-08-420-443-5
34 43.6 2.9 919 2 US-08-924-759-19
35 43.6 2.9 919 3 US-09-248-335-19
36 43.6 2.9 4507 2 US-08-568-459A-3
37 43.6 2.9 4507 2 US-08-487-826B-3
38 43.4 2.9 1931 3 US-09-019-942-2
c 39 43.4 2.9 2502 3 US-09-234-332-1
40 43.2 2.8 2058 2 US-08-749-391-1
41 43.2 2.8 2058 3 US-09-390-200-1
c 42 42.8 2.8 665 2 US-08-883-795A-36
43 42.8 2.8 746 4 US-09-013-810-1
44 42.8 2.8 1736 3 US-09-182-816-22
c 45 42.8 2.8 1736 3 US-09-182-816-24

```

ALIGNMENTS

```

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 23,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9ptc-F1s
US-08-232-463-14

```

```
Query Match          3.7%; Score 57; DB 1; Length 7218;
Best Local Similarity 4.1%; Pred. No. 0.00037;
Matches 15; Conservative 210; Mismatches 140; Indels 0; Gaps 0;

QY 5 CCACCTCTCTCGAATCCCTCCGAGTGGCGCGCTCCAGGCACGCTGCTGAGCAGTCTAC 64
DB 1099 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1158
QY 65 TCCCGCTCACCTCTCTGACATGACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 124
DB 1159 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1218
QY 125 TCTAGAACTCCCTGTTCCAAACCGCTCTCTCGAAACCGTCTGCTGCTGCTGCTGCT 184
DB 1219 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1278
QY 185 ATCTCTTATCTAACCCTCAACACTTCTTCCCGCTTTCATGCAATCTAATCTACCTC 244
DB 1279 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1338
QY 245 TATCGCCGAGAAATGCGGAGTTCGGGTATCAGAACGGTGACTCGGTCTTCTTCAAGA 304
DB 1339 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1398
QY 305 TTATGAGTCTGCTGGAGATCATCGCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 364
DB 1399 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1458
QY 365 GCGAC 369
DB 1459 ACTAC 1463

RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match          3.7%; Score 56.8; DB 1; Length 7218;
Best Local Similarity 4.0%; Pred. No. 0.00041;
Matches 16; Conservative 225; Mismatches 157; Indels 0; Gaps 0;

QY 967 TCGTTTGGGATGGCGAAGATCTCTGCGCGGGATTTGGTCGGAGATGAAGGGGTGTTTCGG 1026
DB 1462 TAGTTAAAGATAGAGAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1403
QY 1027 GCAGCTGAGGCGATCGCGCGGAAATAGAGAAGAGAGAGCGACAGCAAGAAGATTCTAGAA 1086
DB 1402 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1343
QY 1087 ACTGTGGAGAACTGCCCTCTGAGATTCGGAAGCCTTGCAAACTGTATTCTCGGTG 1146
DB 1342 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1283
QY 1147 GCGGATCGACGAGCTTTGATCTTTACGCGCGGATTTTGGATGGGTAAAGCGGTGAAG 1206
DB 1282 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1223
QY 1207 CAAGAGATCTGCTGATTTGAGAGAGAAGTTTACGATGCTGCTGTAAACCGAGGAT 1266
DB 1222 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1163
QY 1267 GCTGCCGCGGATTCGAGGTTGGATTGCTTTGCCAAAGGAGGATTCGAAGCTTTTGTAT 1326
DB 1162 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1103
QY 1327 GATTATTTTGGGAGGAAATAAGGGTTGATTATCAT 1364
DB 1102 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRAT 1065

RESULT 3
PCT-US96-10618-1
; Sequence 1, Application PC/TUS9610618
; GENERAL INFORMATION:
; APPLICANT: Coleman, Roger
; APPLICANT: Guegler, Karl J.
; APPLICANT: Au-Young, Janice
; APPLICANT: Bandman, Olga
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: A NOVEL HUMAN EDG-2 RECEPTOR HOMOLOG
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10618
; FILING DATE: 20-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/000,352
; FILING DATE: 20-JUN-1995
```

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/567,817
; FILING DATE: 06-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Glaister, Debra J.
; REGISTRATION NUMBER: 33,888
; REFERENCE/DOCKET NUMBER: PF-0042 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1875 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; IMMEDIATE SOURCE:
; LIBRARY: Rheumatoid Synovium
; CLONE: 80853
; PCT-US96-10618-1

Query Match 3.5%; Score 52.6; DB 5; Length 1875;
Best Local Similarity 56.7%; Pred. No. 0.0024;
Matches 97; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 1319 CTTTTCATGATTTATTTGGCGGAGGAATAAAGGTTGATTAATCATATTAATCATGATTATTA 1378
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1703 CTTTAAAAAGATTTTGTGTGGTTTGGTGAAGTCAGAATAAATTCGTGCTAGTTGAATCC 1762
QY 1379 TGAAGTTGGATGAATCCCTGCTGTTTCATCTCTATTCTTTAAACAATAATTTTTCATT 1438
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1763 ACNACTTCATTATATACAGGCTCCCTTTTATTTTAAAGGATACGTTTCACCTTAAT 1822
QY 1439 GAACCTTTTGTAGTCAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1489
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1823 AAACACGTTTATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAAT 1873

RESULT 4
US-07-867-106-4/c
; Sequence 4, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostellium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; APPLICATION NUMBER: PCT/AU90/00530
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo

; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3138 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; US-07-867-106-4

Query Match 3.4%; Score 51.4; DB 1; Length 3138;
Best Local Similarity 59.1%; Pred. No. 0.0054;
Matches 88; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
QY 1372 TGTATTATGAAGTTGGATGAATCCCTCTGTTTCATCTCTATTCTTTTAAACAATAATTTT 1431
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2066 TTTATTTTATTTAATTTTATTTGTTTAAAGAAATAAGAAAAAATAAAAAAATAAAAAA 2007
QY 1432 TTCATTGAACCTTTTGTGATCAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1491
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2006 ATCTTTTATGCAATCTGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1947
QY 1492 CAGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1520
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1946 AAAATTTTGAATCCCATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1918

RESULT 5
US-08-242-677-1
; Sequence 1, Application US/08242677
; Patent No. 5677143
; GENERAL INFORMATION:
; APPLICANT: Gaynor, Richard B
; APPLICANT: Wu, Eoon W.
; TITLE OF INVENTION: Cellular Nucleic Acid Binding Protein
; TITLE OF INVENTION: and Uses Thereof in regulating Gene Expression and in the
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/242,677
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mayfield, Denise L.
; REGISTRATION NUMBER: 33,732
; REFERENCE/DOCKET NUMBER: UTSD:401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-787-1400
; TELEFAX: 713-789-2679
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5173 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear


```
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526rls
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; APPLICATION NUMBER: PCT/AU90/00530
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 5852 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
;
; US-07-867-106-2
;
; Query Match
; Best Local Similarity 69.8%; Score 49.6; DB 1; Length 5852;
; Matches 67; Conservative 0; Mismatches 29; Indels 0; Gaps 0;
;
; QY 1427 TTTTTCATTCAGTCTTTTTCAGTCAATAAAAAAAAAAAAAAAAAAAAAATGAAA 1486
; Db 2032 TTTTTCATTCAGTCTTTTTCAGTCAATAAAAAAAAAAAAAAAAAAAAAATGAAA 1973
;
; QY 1487 AACTCAGTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1522
; Db 1972 AAAGTTGGTAAACTACATTAGTTTTTTTTTATAGTTTT 1937
;
; RESULT 10
; US-09-416-050A-1
; Sequence 1, Application US/09416050A
; Patent No. 6194559
; GENERAL INFORMATION:
; APPLICANT: KIM, Soo Young
; TITLE OF INVENTION: Abscisic Acid Responsive Element -Binding Transcription Factors
; FILE REFERENCE: 1942/42
; CURRENT APPLICATION NUMBER: US/09/416,050A
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 83
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-416-050A-1
;
; Query Match
; Best Local Similarity 67.6%; Score 49.2; DB 4; Length 1578;
; Matches 69; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
;
; QY 1388 ATGAAATCCTCTCTTTTCATCTCTATTGTTTAAACAATAATTTTTTCCCATTTGAACCTTTT 1447
; Db 1450 atgaattcataatttcataaacatagttaaaactttcagatttagctatagagaagatg 1509
;
; QY 1448 TGAGTCAATAAAAAAAAAAAAAAAAAAAAAAAAAAATGAAAAA 1489
; Db 1510 tgactaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1551
;
; RESULT 11
; US-09-664-800-1
; Sequence 1, Application US/09664800
; Patent No. 6218527
; GENERAL INFORMATION:
; APPLICANT: KIM, Soo Young
; TITLE OF INVENTION: Abscisic Acid Responsive Element -Binding Transcription Fac
; FILE REFERENCE: 1942/42
; CURRENT APPLICATION NUMBER: US/09/664,800
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/416,050
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-664-800-1
;
; Query Match
; Best Local Similarity 67.6%; Score 49.2; DB 4; Length 1578;
; Matches 69; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
;
; QY 1388 ATGAAATCCTCTCTTTTCATCTCTATTGTTTAAACAATAATTTTTTCCCATTTGAACCTTTT 1447
; Db 1450 atgaattcataatttcataaacatagttaaaactttcagatttagctatagagaagatg 1509
;
; QY 1448 TGAGTCAATAAAAAAAAAAAAAAAAAAAAAAAAAAATGAAAAA 1489
; Db 1510 tgactaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1551
;
; RESULT 12
; US-09-665-309-1
; Sequence 1, Application US/09665309
; Patent No. 6232461
; GENERAL INFORMATION:
; APPLICANT: KIM, Soo Young
; TITLE OF INVENTION: Abscisic Acid Responsive Element -Binding Transcription Fac
; FILE REFERENCE: 1942/42
; CURRENT APPLICATION NUMBER: US/09/665,309
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/416,050
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:01:00 ; Search time 5816.79 Seconds
(without alignments)
2473.399 Million cell updates/sec

Title: US-08-894-356C-6
Perfect score: 1522
Sequence: 1 NTGACCACCCCTCCTCGAATC.....TTTTTTTTTTTTTTTTTTTT 1522

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues 20456230
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est33:*
- 26: gb_est34:*
- 27: gb_est35:*
- 28: gb_est36:*
- 29: gb_est37:*
- 30: gb_est38:*
- 31: gb_est39:*
- 32: gb_est40:*
- 33: em_estba:*
- 34: em_estfun:*
- 35: em_esthum1:*
- 36: em_esthum2:*
- 37: em_esthum3:*
- 38: em_esthum4:*
- 39: em_esthum5:*
- 40: em_esthum6:*
- 41: em_esthum7:*
- 42: em_esthum8:*
- 43: em_esthum9:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estp11:*
- 73: em_estp12:*
- 74: em_estp13:*
- 75: em_estp14:*
- 76: em_estp15:*
- 77: em_estp16:*
- 78: em_estp17:*
- 79: em_estp18:*
- 80: em_estp19:*
- 81: em_estp110:*
- 82: em_estro1:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estro10:*
- 92: em_estro11:*
- 93: em_estro12:*
- 94: em_estro13:*
- 95: em_estro14:*
- 96: em_estro15:*
- 97: em_estro16:*
- 98: em_estro17:*
- 99: em_estro18:*
- 100: em_estro19:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
```

```
190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	159	10.4	763	155	BG591758	EST499600
2	157.2	10.3	690	155	BG599447	EST504342
3	146.6	9.6	606	113	AW221049	EST297518
4	138.8	9.1	570	142	BE922784	EST426553
5	138.6	9.1	583	113	AW221050	EST297519
6	124.8	8.2	564	142	BE921494	EST425179
7	123	8.1	574	142	BE922746	EST426515
8	121.8	8.0	687	111	AW102336	sd86d06.7
9	119.4	7.8	508	173	BG097054	EST461573
10	116	7.6	591	167	BE434257	EST405335
11	114.8	7.5	659	155	BG598154	EST496832
12	112	7.4	518	119	AW650280	EST328734
13	108	7.1	487	118	AW616206	EST307245
14	106.8	7.0	482	144	BF096979	EST400516
15	97.8	6.4	456	142	BE923572	EST427341
16	97	6.4	386	173	BG097680	EST462199
17	97	6.4	411	142	BE922378	EST426147
18	97	6.4	414	142	BE921952	EST425721
19	95.4	6.3	368	142	BE921953	EST425722
20	92	6.0	307	142	BE920010	EST423779
21	86.4	5.7	497	122	AW930934	EST356777
22	86	5.7	408	142	BE920022	EST423791
23	85.8	5.6	354	142	BE922950	EST426719
24	85.4	5.6	403	143	BF052576	EST437806
25	85.2	5.6	783	155	BG584708	EST486469
26	84.8	5.6	490	110	AW038515	EST280198
27	83.2	5.5	409	111	AW093479	EST286659
28	81.8	5.4	575	142	BE920468	EST424237
29	81.8	5.4	788	164	BE187619	EST336180
30	79.2	5.2	597	120	AW774948	EST334099
31	79	5.2	654	151	BF642724	NF070E031
32	79	5.2	656	119	AW692527	NF056D05S
33	79	5.2	677	118	AW586972	EST318595
34	78.8	5.2	708	138	BE660849	7-c8 Gmax
35	77.2	5.1	563	152	BG362734	sac08a08.
36	77	5.1	311	166	BE353747	EST355090
37	76.6	5.0	334	143	BF053201	EST438431
c 38	73.4	4.8	605	30	AV440958	AV440958
c 39	71.6	4.7	1101	219	CNS00293	Drosophila
40	71.2	4.7	676	153	BG455847	NF071A12P
41	71.2	4.7	681	151	BF639071	NF079A08P
42	70.2	4.6	526	166	BE323055	NF001A07P
43	70	4.6	422	155	BG589694	EST497536
44	69.8	4.6	649	151	BF643475	NF004H05E
45	69.8	4.6	694	117	AW560686	EST315734

ALIGNMENTS

RESULT	1
BG591758	
LOCUS	BG591758
DEFINITION	EST499600 P. infestans-challenged leaf Solanum tuberosum cDNA clone
ACCESSION	BPL110A14 5' sequence, mRNA sequence.
VERSION	BG591758
KEYWORDS	EST.
SOURCE	EST.
ORGANISM	potato.
REFERENCE	Solanum tuberosum.
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 763)
Zhang, P., Hernandez, M., Tornqvist, C.-E., Wirtz, U., Loukolanov, A., Rangel, P., Haberland, G.T., Cho, J., Chiemingo, A., Bougri, O., Buell, C.R., Ronning, C.M., Helgeson, J. and Baker, B.

TITLE Generation of ESTs from Potato Leaves Challenged with Phytophthora infestans, Incompatible Reaction
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.

FEATURES

source
1..763
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="BPL110A14"
/clone_lib="P. infestans-challenged leaf"
/tissue_type="leaf"
/dev_stage="6 week old"
/lab_host="SOLR"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Whole plants were challenged with 450,000 sporangia/ml P. infestans US-1(US 940501) in Biotron (Madison, Wisconsin). Leaf tissue was collected at 1, 2, 5, 12, and 24 hours post-challenge and frozen in liquid nitrogen immediately upon removal. Kennebec plants showed no signs of HR. Katabdin plants (susceptible to P. infestans US-1) were used as controls and showed infection. NOTE: We cannot exclude the possibility that this sequence is actually derived from Phytophthora rather than potato."

BASE COUNT 212 a 162 c 156 g 233 t
ORIGIN

Query Match 10.4%; Score 159; DB 155; Length 763;
Best Local Similarity 56.6%; Pred. No. 7.2e-23;
Matches 314; Conservative 0; Mismatches 240; Indels 1; Gaps 1;

QY	25	CGAGTGGCGCGCTCCAGCC-ACGGTGGCTGAGCAGTCACTCCCGCTCACTTCTTCGA	83
DB	4	CAAGTGGCGCCACTCGGCGGGCCGCGGAGGTGACACTCCCTTACTTATTGA	63
QY	84	CATGAGTGGCTGCATTTCCACCCATGCTTACGTTCTCTTACGAACTCCCTTC	143
DB	64	TCATGTTGGTTAGGTTCCACGTATACGCGGATATTCTACAGCTCCCAATTC	123
QY	144	CAACCCGCTTCTCGAACCCTGTTCCGAACTCAACAATCTTATCTAACCT	203
DB	124	CAACCCGATTCGTTCAAAACATTTCTACTCTTTAAATAATTCACCTCTCCCTCACT	183
QY	204	CAACACTTCTTCCCTTTTCATGCAATCTAATCTACCCCTCTATCGCGGAGAAATGCC	263
DB	184	CAACACTACACCCCTTAGCAGGAATGTTGCTCCACTAGATACAAATGGATATCC	243
QY	264	GGAGTCCGGTATAGAACCGTGACTCGGTTCTTTTCACGATTATGGAGTCTGTCGAGA	323
DB	244	TGAGTACGTTATGTACAGGAAATCTGTATCTGTACTTTTTTTTGAGATATGGA	303
QY	324	TCATCCGCTTCCGCTCATAAATACTACTCTTTTGGCCCTAGCGAGATATGAAGATCT	383
DB	304	TTTCAATATCTCATTTGGTGACCATCGCGTAAATGCTAAGGATTTTATCATTCTTGTTC	363
QY	384	CCAGCTGCCCGCATAGTCGAGGAATCTGATCGGAATTTCTTCAAGTTTTTCCCGTGCA	443
DB	364	TAAGTAGGGGAACCTAAGGATGCACCCGGGTCCCAATTAGCTCCGCTTAGCCATTCA	423
QY	444	AGTGACTCTGTTTCCGGTTCGCGGGTGTGCATCGGAATAACGACGACACACCGTTAG	503
DB	424	AGTGACACTTTTTCGAAATCTTGTGTATCCATTTGTTTTCACATAACCATCATGTGTGG	483
QY	504	CGATGCTCCATCTGTTGTAGGGTTTATGAAGAGTTGGGCTTCCATCACTAAATTCGGAG	563
DB	484	TGATGGAGCTACCATAGTAGGGTTTCAATAGGCGTGGGCTCTACTCCATAAATTCGGTGG	543

```
Qy 564 AGATGATGAATTCCT 578
Db 544 AGATGAACAATTCCT 558

RESULT 2
BG599447 690 bp mRNA EST 12-APR-2001
LOCUS EST504342 cSTS Solanum tuberosum cDNA clone cSTS25122 5' sequence,
DEFINITION mRNA sequence.
ACCESSION BG599447
VERSION BG599447.1 GI:13616583
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 690)
Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13P-R.
FEATURES
source
1..690
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS25122"
/clone_lib="cSTS"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
BASE COUNT 178 a 156 c 144 g 212 t
ORIGIN

Query Match 10.3%; Score 157.2; DB 155; Length 690;
Best Local Similarity 55.2%; Pred. No. 1.7e-22;
Matches 306; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

Qy 25 CGAGTGGCGCCCTCCAGGACAGGTGGCTGAGAGTACACTCCGCTCCTTCGAC 84
Db 9 CAAGTGGCGCCACCTCCGCGCGCGACGAGGTGACACTCCCTCTACTATTATTTGAT 68
Qy 85 ATGAGTGGCTGCTATTCACCCCATGCTTCAGCTTCCTCTACGAACTCCCTGTTC 144
Db 69 CATGTTGGTTAGGTTCCACCGTATAGCGGGATATATTCTACAAGTCCCATTTCC 128
Qy 145 AAACCCGCTTCCTCGAAACCGTCTGTCGGAACCTCAACAACTCTTATCTTAACCTC 204
Db 129 AAACCCGATTTTCGTTCAAAACATATTCTCTCTTAAAAATTCACCTCTCCCTCACTC 188
Qy 205 AAACACTTCTCCCTTCATGCAATCTATCTACCTCTATCGCCGGAGAAAATGCCG 264
Db 189 AAACACTATAGCCCTTAGCTGGAAACGTGCTCGTCCACTAGATACAACGGATATCCT 248
Qy 265 GAGTTCCGGTATCAGACGGTGACTCGGTTTCTTCAGGATATTGGAGTCGTGGAGAT 324
Db 249 GAGTTACATTATGTGACAGGAGATTCGTATCTGTACTCTGTTCTGAGACATGATGAT 308
Qy 325 CATCCGCATCCGGCTCAATAATACTACTGTCTTTCGCCCTAGCGACGATTTATGAAGATCTC 384

Db 309 TTCAATTATCTCATTTGGTGACCATCCGCGTAATGCTAAGGATTTTATCACTTTGTTCT 368
Qy 385 CAGTCGCGCGGATAGTCGAGGAATCTGATCGGAAATGTTTCAAGTTTTCAGCGGTGCAA 444
Db 369 ACGTTAGGGGAACCTAAGGATGCACCCGGGTCCAAATAGCTCCGGCTTATGACCAATCAA 428
Qy 445 GTGACTCTGTTTCCCGGTCGCGGGGTGTCATCGGAATAAGCAGCACACACCGCTTAGC 504
Db 429 GTGACACTTTTCCGATCTTGGTGATCCATGTTGTTTCACTAACCATCATGTTGTTGGT 488
Qy 505 GATGCTCCATCGTTTGTAGGGTTTATGAAGAGTTGGGCTTCCATCACTAAATTCGGAGGA 564
Db 489 GATGGAGCTACCATAGTAGGTTTCATTAGGGCGTGGGCTCTACTCCATAAATTCGGTGA 548
Qy 565 GATGATGAATTCCT 578
Db 549 GATGAACAATTCCT 562

RESULT 3
AW221049 606 bp mRNA EST 07-DEC-1999
LOCUS EST297518 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION cDNA clone cLEF3f1, mRNA sequence.
ACCESSION AW221049
VERSION AW221049.1 GI:6532733
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 606)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.,
Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
FEATURES
source
Location/Qualifiers
1..606
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF3f1"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
BASE COUNT 147 a 143 c 123 g 193 t
ORIGIN

Query Match 9.6%; Score 146.6; DB 113; Length 606;
Best Local Similarity 53.4%; Pred. No. 2.6e-20;
Matches 308; Conservative 0; Mismatches 289; Indels 0; Gaps 0;

Qy 2 TGACCACCCTCCTCGAATTCCTCCGAGTGGCGCGCTCCAGGCACGGTGGCTGAGCAGT 61
```


TITLE Generation of ESTs from tomato fruit tissue
JOURNAL Unpublished (1999)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU

5 prime sequence.
FEATURES Location/Qualifiers

source
1..583
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF3F3"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLEF - Fruit were tagged at the 1cm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"
BASE COUNT 142 a 140 c 117 g 184 t
ORIGIN

Query Match 9.1%; Score 138.6; DB 113; Length 583;
Best Local Similarity 52.5%; Pred. No. 1.1e-18;
Matches 303; Conservative 0; Mismatches 274; Indels 0; Gaps 0;
Qy 2 TGACCACCTCTCGAATCTCCGAGTGGCGCCCTCCAGGACGGTGGCTGAGCAGT 61
Db 5 TGGCCACCGTGATTGAGCAATGTCAAGTTGGCCACCTCCGGCGGCGCAACGGAGTGA 64
Qy 62 CACTCCCGCTCACCCTCTTCGACATGAGCTGCTGCATTTCCACCCCATGCTTCAGCTTC 121
Db 65 TACTCCCTCTTACTATTTTGACCATGTTGGTTAGGCTTTGCGCGTATGAGCGGATAT 124
Qy 122 TCTTCTAGCAACTCCCTCTTCCAAACCCGCTTCTCGAAACCGTCTCGGAAACTCA 181
Db 125 TATTTTACAAGCTCTCCATTTTCAAAACCGATTGCTTCAAAACATTTATCTCTCTCTTA 184
Qy 182 AACATCTTATCTACCTCAACACTTCTTCCCTTTTCATGCAATCTAATCTACC 241
Db 185 AAAATCACTCTCCCTCACTCTCAACACTATAGCCCTTAGCCGGAACGTTGCTTCTC 244
Qy 242 CTCTATCCCGGAGAAAATGCGGAGTTCGGGTATCAGACGGTCACTCGGTTCTTTTCA 301
Db 245 CACTAGATACAACGGATATCTGATTTACGTTATGTGACAGGAGATTTCTGTCTGTGA 304
Qy 302 CGATTATGAGTCTGTGCGGATCATCCGCAATCCGGTCTATAAATPACTCTCTTGGCCC 361
Db 305 CTTTTCGAGACTGATATGAATTTCAATTTATCTCATTGGTGACCATCCGCGTAAGGCTA 364
Qy 362 CTAGGAGCATATGAAGATCTCCAGCTGCGCGCATAGTGCAGGAATCTCATCGGAAT 421
Db 365 AGGATTTTATCATTGTTCTTAAAGTTAGGGAACCTTAAGGATGCACCGGGGTCTCAAC 424
Qy 422 TGTTCAGATTTTATAGCCGTGCAAGTACTCTGTTTCCGGTTCGGGGTGTGTCGGA 481
Db 425 TAGCCCGCTCTTAGCCATTCAGGTGACACATTTTCCGAATCTTGGTGATCCATGGTT 484
Qy 482 TAACGACACACACACCGTTAGCGATGCTCCATGCTTTTGTAGGGTTTATGAAGAGTTGG 541
Db 485 TCACATAACCATCATGTTCTTGGTGATGGAACACTATATAGCATGTTTCAATTAAGCGTGG 544
Qy 542 CTTCCATCACTAAATTCGGAGAGATGATGATTCCT 578
Db 545 CTCCTACCAAAATTAGTGGACATGAACCAATCTT 581

RESULT 6

BE921494
LOCUS BE921494 564 bp mRNA EST 02-OCT-2000
DEFINITION EST425179 potato leaves and petioles Solanum tuberosum cDNA clone
CSTB13B18 5' sequence, mRNA sequence.
ACCESSION BE921494
VERSION BE921494.1 GI:10447486
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 564)
AUTHORS van der Hoeven, R.S., Bezzerides, J., Holt, I.E., Liang, F., Cho, J.,
Utterback, T., Hansen, C.L., Doan, B., Bougri, O., Buell, C.R., Ronning,
C.M., Fry, W.E., Tanksley, S.D. and Baker, B.
Generation of ESTs from potato leaves and petioles
Unpublished (2000)
CONTACT: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES Location/Qualifiers

source
1..564
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="CSTB13B18"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 150 a 139 c 97 g 178 t
ORIGIN

Query Match 8.2%; Score 124.8; DB 142; Length 564;
Best Local Similarity 53.6%; Pred. No. 7.9e-16;
Matches 283; Conservative 0; Mismatches 242; Indels 3; Gaps 1;

Qy 28 GTGGCGCGCTCCAGGACGGTGGCTGAGCAGTCACTCCGGCTCACCTTCTTCGACATG 87
Db 40 GCGCACCTCCCGCGGCGGCGACAGCTAAGCTCCCTCTAATCTATTTGATCAT 99
Qy 88 ACGTGGCTGCATTTCCACCCCATGCTTCAGCTTCCTTCTAGGAACCTCCCTGTTTCCAA 147
Db 100 ATATGGTTGGTGTGGCTATATGCGGATTTTATTTCTACAAGCTACCAATTTCCAA 159
Qy 148 CCGCCCTTCGAAACCGTCTGTTCCGAACTCAACAACTCTTATCTCTAACCTCAAA 207
Db 160 CTCGATTTCTGTTCAACCAATTTATCTACTCTTAAACATCTCACTCTCCCTCACTCA 219
Qy 208 CACTTCTTCCCTTTCATGCAATCTAATCTACCTCTCTATCGCGGAGAAAATCGCGAG 267
Db 220 CACTACACACCTTAGCTGGCAACATTCCTGTCCCAATTAATTCGAGTGGTATCCTGAG 279
Qy 268 TTCCGGTATCAGAACGGTGAATCGGTTTCTTTCAGGATTTATGAGTCTGTGCGGATCAT 327
Db 280 TTGGTATGATGACTGGAGATTCATATCTGTACTTTTATTTAGAGCATGATGAATTC 339
Qy 328 CCGCATTCGGCTCATAAATACTACTGCTTTGCCCTACGAGGATTTATGAAGATCTCCAG 387
Db 340 AATCATCTCATTTGGTAACCATCTCTCGAAATG---CTAAGGATTTTATCCCATCTCTCA 396

All-Purpose potting soil was supplemented with: 0.36g/L available phosphoric acid (p205), 20mg/L urea N, 0.16g/L S, 0.49mg/L B, 2.5mg/L Cu, 0.15g/L Fe, 13.53mg/L Mn, 0.26mg/L Mo, 14mg/L Zn, 20mg/L Ca, and the following nutrients in a slow-release form (Osmocote): 0.165g/L ammonia N, 0.185g/L nitrate N, 0.35g/L available phosphoric acid, and 0.35g/L soluble potash. No nodules were visible on the roots at harvest. Stratagene's cDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA is hemimethylated.

Stratagene's first-strand synthesis primer was used [GAGAGAGAGAGAGAGAGAGAGAG(T)-18]. After second-strand synthesis is, the cDNA ends were 'polished' with clone Pfu DNA polymerase, ligated to EcoRI adapters, and phosphorylated. The XhoI site within the first-strand synthesis primer was restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 400bp cutoff, using a SizeSeph 400 Spun column from Pharmacia. The column eluent was then ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+)) that had been digested with EcoRI and XhoI, and phosphorylated. Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts. This library was constructed by Dr. Paul Kelm and Dr. Virginia Coryell."

BASE COUNT 157 a 247 c 91 g 187 t 5 others
ORIGIN

Query Match 8.0%; Score 121.8; DB 111; Length 687;
Best Local Similarity 52.7%; Pred. No. 3.3e-15;
Matches 286; Conservative 0; Mismatches 254; Indels 3; Gaps 1;

```

QY 27 AGTGGCGGCTTCAGGACGGTGGCTGAGGAGTCACTCCGCTCAGCTTCTTCGACAT 86
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 81 AGTTGGTCCACCACATGCTCACTTCCCTCAACCACTTCCCTTCTTCGATAT 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 87 GACGTGGTGCATTTCCACCCCATGCTTCAGCTTCTCTTACGAACTCCCTGTTCAC 146
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 141 TCCATGGTCTACTGCCACCCCATCCAGGATCTTCTCTATGACTTCCGCCACCCAC 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 147 ACCCGCTTCCTCGAAACGGTGGTTCGAAACTCAAACTCTTATCTCTAACCTCAA 206
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 201 ACACCACTTCTCCAAACAGACTTCCCATCTCAACACACTCTTCTCCCTAACCTCCA 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 207 ACACCTTCTCCCTTTCATGCAATCTAATCTACCTCTATCGCGGAGAAATCCCGGA 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 261 ACACCTTCTCCCTTCTCTCCAAATCTCATCTGTTCTCCACAGCAACCCCATCTCTCA 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 267 GTTCCGGTATCAGACGGTGAAGTCTGTTTTCACGATTTATGGAGTCTGTCCGAGA 323
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 321 CATAGCTACCTTGATGGAGACTCTCTCTCTTACCGTTGAGAGTCCACCGCACTT 380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 324 TCATCCGATTCGCTCATAAATACTACTGCTTTTCCCTTACGCAAGTATATGAAGATCT 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 381 CACCTCTCTACATCAGATTTCACCAAGAGCTTCCAAATTTGGCACCTCTGTTCGCGC 440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 384 CCAGCTGGCGGATAGTCGAGGAATCTGATCGGAAATTTGTTCAAGTTTATGCGGTGCA 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 441 GTTTCCTACCCACAGTGTGACCAAGATGGCGCAGTGTGTCTCTTATGGCCATTTCA 500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 444 AGTGACTCTGTTCCGCTGCGGGGTGTGATCGGAATACGAGGCACACCGTTAG 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 501 GGTACATTTTCCCAAGACTGGCTTCACCATATGTCTCACCTTCACACCTTGCAG 560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 504 CGATGCTCCATGTTGTAGGGTTTATGAAGAGTTGGCTTCCATCACTAAATTCGGAGG 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 561 CGACGGCAAGTCACTTCACCAATTTTCATCAAGTTCCTGGGCTCTCTTTCANAGCANAAG 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 564 AGA 566
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 621 AAA 623

RESULT 9

LOCUS BG097054 508 bp mRNA EST 29-JAN-2001
DEFINITION EST461573 potato leaves and petioles Solanum tuberosum cDNA clone
cSTB45p9 5' sequence, mRNA sequence.
ACCESSION BG097054
VERSION BG097054.1 GI:12587089
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 508)
AUTHORS van der Hoeven, R.S., Bezzerides, J., Holt, I.E., Liang, F., Cho, J.,
Utterback, T., Hansen, C.L., Doan, B., Bougri, O., Buell, C.R., Ronning,
C.M., Fry, W.E., Tanksley, S.D. and Baker, B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
Location/Qualifiers
1..508
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB45p9"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."

BASE COUNT 127 a 130 c 87 g 164 t
ORIGIN

Query Match 7.8%; Score 119.4; DB 173; Length 508;
Best Local Similarity 53.0%; Pred. No. 1e-14;
Matches 255; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

```

QY 28 GTGGCGCGCTCCAGGACCGTGGCTGAGCAGTCACTCCCGCTCACCTTCTTCGACATG 87
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 17 CGCCACCTCCCGCGCGGCGGACAGAGCTACGCTCCCTCTAACTTATTTGATCAT 76
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 88 ACGTGGCTGCAATTTCCACCCCATGCTTCAGCTTCTTCTTACGAACCTCCCTGTTCACAA 147
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 77 ATATGTTTCGTTTGGCTATATCGTCGGATTTTATTCTACAAGCTACCAATTTCCAAA 136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 148 CCGGCTTCCTCGAAACCGTGTTCGGAACCTCAACAACTCTTATCTCTAACCTCAAA 207
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 137 CTCGATTTCTGTTCAACCAATTTATCTACTCTTAAACACTTCACTCTCCCTCACTCAA 196
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 208 CACTTCTTCCCGCTTTCATGCAATCTAATCTACCTCTATCGCGGAGAAAAATCGCGAG 267
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 197 CACTACACACCTTAGCTGGCAACATTCCTTGTCCACTAAATTCGAGTGGTATCCTGAG 256
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 268 TTCGGTATCAGACGGTGAATCGGTTTCTTTCACGATTTGAGTCTGTTCGGAGCATCAT 327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 257 TTGCGTTATGTGACTGGAGATTCATATCTGTACTTCTTATTTAGACTGATGAATTC 316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 328 CCGCATTCGGCTCAATAATACTACTGCTTTGCCCTACGAGCATTTATGAAGATCTCCAG 387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 317 AATCATCTCATTTGGTAACCAATCTTCGAAATGCTAAGGATTTTATCCCTTCAATCCCTCAA 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
QY 388 CTGCCCCCGGATCTCGAGGAATCTGATCGGAATTTCTTTCAAGTTTATGCCGGTGAAGTG 447
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 TTGGCAACACCTAAGGATGCACGGGGTCAAACTAGTCCGGCTCTAGCCATTTCAAGTG 436
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 448 ACTCTCTTTTCCGGTCCGGGGTGTGCATCGGAATAACGACGACACACCGCTTACCGAT 507
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 437 ACATTTTTCGGAATCTTGGCATATCCGTTGGTTTGTAGTAACCATCATGTCGCTTGTGAT 496
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 508 G 508
      |
Db 497 G 497

RESULT 10
LOCUS BE434257
DEFINITION EST405335 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
      clone cLEG15H6, mRNA sequence.
ACCESSION BE434257
VERSION BE434257.1 GI:9432100
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
      Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
      Lycopersicon.
REFERENCE 1 (bases 1 to 591)
AUTHORS Alcalá,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
      Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
      Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
      ,S.D.
      Generation of ESTs from tomato fruit tissue, breaker stage
      Unpublished (2000)
      Contact: David Frisch
      Clemson University Genomics Institute
      Clemson University
      100 Jordan Hall, Clemson, SC 29634, USA
      Tel: 864 656 4366
      Fax: 864 656 4293
      Email: dfrisch@clemson.edu
      5 prime sequence.
FEATURES
      source
      1..591
      /organism="Lycopersicon esculentum"
      /cultivar="TA496"
      /db_xref="taxon:4081"
      /clone="cLEG15H6"
      /clone_lib="tomato breaker fruit, TIGR"
      /tissue_type="Pericarp"
      /dev_stage="breaker"
      /lab_host="SOLR"
      /note="Vector: pBluescriptSKmCuadapt; Site_1: EcoRI;
      Site_2: XhoI; Fruit were harvested at the breaker stage
      (first sign of lycopen accumulation on the blossom end of
      the fruit). Fruit were cut in half and the seeds and
      locules were discarded prior to freezing the pericarp."
BASE COUNT 159 a 129 c 121 g 182 t
ORIGIN

Query Match 7.6%; Score 116; DB 167; Length 591;
Best Local Similarity 52.7%; Pred. No. 5.1e-14;
Matches 251; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

QY 103 CACCCCATGCTTCAGCTTCTCTTCTACGAACTCCCTGTTTCCAAACCGCGCTTCCTCGAA 162
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 CGCCGTATGAGCGGATATATTATTTACAAGCTCTCCATTTTCAAAACCGGATTTCTGTTCAA 60
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163 ACCGTGTTCCGAACTCAACAATCCTTATCTTAACCTCAACACACTCTCTCCCCCTT 222
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AACATTATTCCTCTTAAATAATCACTCTCCCTCACTCTCAACACACTATACGCCCTTA 120
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 223 TCATGCAATCTAATCTACCTCTATCGCCGGAGAAAATCCGGAGTTCGGGTATCAGAAC 282
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GCCGAAACGTTGCTTGTCTCAGCTAGATACAAACGGATATCTCTGAGTTAGTTATGTGACA 180
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 283 GGTGACTCGGTTCTTTTCAGGATTATGGAGTCTGTGCGGAGATCATCCGATCCCGCTCAT 342
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 GGAGATCTGTGTCTGTTACTTTTTTCGAGACTGATATGAATTTCAATTTATCTCATTTGGT 240
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 343 AAATACTACTGCTTTTGCCTTAGCGACGATTATGAAGATCTCCAGCTCGCGCGGATAGTC 402
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 GACCATCCCGTTAAGCTTAAGGATTTTTTATCACTTTGTTCTTAAGTTAGGGNAACCTAAG 300
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 403 GAGGAATCTGATCGGAAATTTGTTTCAAGTTTATAGCGGTCAAGTACTCTGTTTCCCGGT 462
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 GATGCACCGGGGTCCAACTAGCCCGCTCTTAGCCATTCAGGTGACACTTTTTCGGAAT 360
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 463 CGCGGGTGTGCATCGGAATAACGACGACACACCGCTTAGCGATGCTCCATGCTTTGTA 522
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 CTGGGTGTTATCCATGGTTTCACTAACCATCATGTTGTTGGTGATGGAGCTACTATAGCA 420
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 523 GGGTTTATGAAGAGTTGGGTTCCTCATCACTAAATTCGAGGAGATGATGAATTCCTT 578
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 GGGTTTATGAAGCGTGGGTCTACTCCACAAATTCGGTGGACATGAACAATTCCTT 476
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
LOCUS BG598154
DEFINITION EST496832 cSTS Solanum tuberosum cDNA clone cSTS20M23 5' sequence,
      mRNA sequence.
ACCESSION BG598154
VERSION BG598154.1 GI:13616294
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
      Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 659)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
      Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
      Generations of ESTs from sprouting potato eyes
      Unpublished (2000)
      Contact: Cathy Ronning
      The Institute for Genomic Research
      For clone info: please contact Research Genetics, Libraries
      Division tel 1-800-711-6195, email cdnaeresgen.com
      Seq primer: M13F-R.
FEATURES
      source
      1..659
      /organism="Solanum tuberosum"
      /cultivar="Kennebec"
      /db_xref="taxon:4113"
      /clone="cSTS20M23"
      /clone_lib="cSTS"
      /tissue_type="sprouting eyes from tubers"
      /dev_stage="12-14 weeks post harvest"
      /lab_host="SOLR"
      /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
      XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
      taken from tubers. The tubers were incubated at 26C in the
      dark for 2-3 weeks prior to sprouting. The eyes were
      frozen in liquid nitrogen immediately upon removal from
      tubers."
BASE COUNT 178 a 163 c 112 g 206 t
ORIGIN

Query Match 7.5%; Score 114.8; DB 155; Length 659;
Best Local Similarity 52.9%; Pred. No. 9e-14;
Matches 293; Conservative 0; Mismatches 257; Indels 4; Gaps 2;
```

```
Qy 28 GTGGCGCGCTCCAGGACCGGTGGCTGAGCAGTCACTCCCGCTCACTCTTCTTGACATG 87
Db 107 GCGCCACTCCCGCCAGCGCGGCAACAGAGCTAACGCTCCCTCTAACTATTTTGATCT- 165
Qy 88 ACGTGGCTGATATTTCCACCCCATGCTTCAGCTTCTCTCTACGACCTCCCTGTTCACAA 147
Db 166 ATATGGTTGGTGGTGGCTATATCGCTGGGATTTTATTTCTACAAGCTACCAATTTCCAAA 225
Qy 148 CCGCGCTTCCTCGAAACCGTGGTTCGGAACCTCAACAACTCTTATCTCTAACCTCAAA 207
Db 226 CTCGATTCGTTCAACACCATTTATCTCTACTCTTAACATTCACCTCTCCCTCACTCAAA 285
Qy 208 CACTTCTTCCTCCCTTCATGCAATCTAATCTACCTCTATCCGCGGAGAAAATCCCGAG 267
Db 286 CACTACACACCTTAGCTGGCAACATGCTTGTCCACTAAATTCGAGTGGTTATCCTGAG 345
Qy 268 TTCGGTATCAGACGGTGACTCGGTTCTTCTTCAGGATTTAGGAGTCTGCGGAGATCAT 327
Db 346 TTGGTTATGTGACTGGAGATTTCTATCTCTTCTTCTTATTTAGACTGATGATGAATTC 405
Qy 328 CCGCATTCGGCTCATAAATACTACTGCTTTCGCCCTAGCGACGATTTATGAAGATCTCCAG 387
Db 406 AATCATCTCATTTGGTAACCATCTCGAATGCTAAGGATTTTATCCCTTCATTCCTCAA 465
Qy 388 CTGGCGCGGATGTCGAGGAATCTGATCGGAAATTTGTTTCAAGTTTATGCGGTGCAAGTG 447
Db 466 TTGGCACAACTAAGGATGACCGCGGGGTCAAACTAGTCCCGGCTTAGCCATTCAGAGT 525
Qy 448 ACTCTGTTTCCCGTCCGCGGGGTGT---GCATCGGAATAACGACGACACACCGTTAGC 504
Db 526 ACACCTTTTCCGAATCTTGGCATATCCCGTTGGTGTTTAGTAAACCCATCATGTGCGCTGT 585
Qy 505 GATGCTCCATCGTTGTAGGGCTTTATGAAGAGTTGGGCTTCCATCACTAAATTCGGAGGA 564
Db 586 GATGGAATACCATCGTGAATTCATAGAACATGGGGTTTACTCAACAATTCGCGGT 645
Qy 565 GATGATGAATTCCT 578
Db 646 GATGAACAGTCTT 659
```

RESULT 12

```
AW650280
LOCUS EST 518 bp mRNA 04-APR-2000
DEFINITION EST328734 tomato germinating seedlings, TAMU Lycopersicon
esculentum cDNA clone cLEI12F13 5', mRNA sequence.
ACCESSION AW650280
VERSION AW650280.1 GI:7411518
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 518)
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Giovannoni,J.J., Martin,G.B. and Tanksley
,S.D.
```

TITLE Generation of ESTs from germinating tomato seed
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.

FEATURES
source Location/Qualifiers
1..518 /organism="Lycopersicon esculentum"

```
BASE COUNT 134 a 119 c 100 g 165 t
ORIGIN
/cultivar="TA96"
/db_xref="taxon:4081"
/clone="cLEI12F13"
/clone_lib="tomato germinating seedlings, TAMU"
/tissue_type="whole seedlings"
/dev_stage="7 days post imbibition"
/notes="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; 7 days post imbibition on water-agar. Mixed stage
whole germinating seedlings from seed coat emergence up
to two centimeters in seeds not showing obvious signs of
germination were discarded."
```

Query Match 7.4%; Score 112; DB 119; Length 518;
Best Local Similarity 51.8%; Pred. No. 3.4e-13;
Matches 253; Conservative 0; Mismatches 235; Indels 0; Gaps 0;

Qy 91 TGGCTGCATTTCCACCCCATGCTTCAGCTTCTCTTACGAACCTCCCGCTGTTCACAAACCC 150

Db 4 TGGTTAGGTTTCGCGGTATGAGCGGATATATTTACAAGCTCTCCATATCAACCC 63

Qy 151 GCCTTCCTCGAAACCGTTCGCGAACTCAACAACTCTTATCTCTAACCTCAACAC 210

Db 64 GATTTCTCTCAAAACATATTTCTCTCTCTTAAAAAATTCACCTCTACCTCAACAC 123

Qy 211 TTCTTCCCTTTCATGCAATCTAATCTACCTCTATCGCGGAGAAAATGCCGAGTTC 270

Db 124 TATACGCCCTTAGCCGGAAGCTTGTGTCACATAGATACAAACGGATATCTCGAGTTA 183

Qy 271 CGGTATCAGAACGGTGACTCGGTTCTTTTACGATATGAGTCTGTGAGAGATCATCCG 330

Db 184 CTTTATGTGACAGAGATGCTGAGTCTGTACTTTTTCGAGACTGATATCAATCTCAAT 243

Qy 331 CATTCGCTCATAAATACTACTGCTTTGCCCTTAGCGGCTAAGGATTTTGTATCACTTT 390

Db 244 TATCTCATTTGGTGACCATCGCGCTAAGGCTAAGGATTTTGTATCACTTTTGTCTT 303

Qy 391 CCGCGGATAGTCGAGGAATCTGATCGGAAATTTTCAAGTTTTCGCGTCAAGTGACT 450

Db 304 GGGAACTTACGATGACCGGGGTCCACTAGCCCGCTCTTAGCCATTCAGTGACA 363

Qy 451 CTGTTTCCCGTCCGCGGGGTGTGATCGGAATAACGACGACACACCGTTAGCGATGCT 510

Db 364 CTATCTCCGAATCTTGGTGATCCATTTGATTTCACTAACCATCATGTTGTGGTGATGA 423

Qy 511 CCATCGTTTGTAGGGTTTATGAAGATTTGGCTTCCATCACTAAATTCGGAGGAGATGAT 570

Db 424 GCTACTATAACAGGGTTTCATTAAGCGCTGGGCTCTACTCCACAAATTCGGTGGACATGA 483

Qy 571 GAATTCCT 578

Db 484 CAATTCCT 491

RESULT 13

```
AW616206
LOCUS EST 487 bp mRNA EST 24-MAR-2000
DEFINITION EST307245 L. hirsutum trichome, Cornell University Lycopersicon
hirsutum cDNA clone cLHT1D15 5', mRNA sequence.
ACCESSION AW616206
VERSION AW616206.1 GI:7322240
KEYWORDS EST.
SOURCE Lycopersicon hirsutum.
```

ORGANISM Lycopersicon hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.

REFERENCE 1 (bases 1 to 487)
AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang
,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W.,

TITLE		Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D. Generation of ESTs from wild tomato (<i>Lycopersicon hirsutum</i>) trichomes					
JOURNAL COMMENT		Unpublished (2000) Contact: David Frisch Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrische@CLEMSON.EDU 5 prime sequence.					
FEATURES source		Location/Qualifiers 1..487 /organism="Lycopersicon hirsutum" /db_xref="taxon:62890" /cloned="clHTD15" /clone_lib="L_hirsutum trichome, Cornell University" /tissue_type="trichome" /dev_stage="mixed stages" /note="Leaves of various stages were shaken in liquid nitrogen, shearing off trichomes. This procedure yielded a mixture of cells which is highly enriched for trichome likely with minor contaminations of other types of leaf cells."					
BASE COUNT		116 a 117 c 91 g 163 t					
ORIGIN							
Query Match		7.18; Score 108; DB 118; Length 487;					
Best Local Similarity		51.7%; Pred. No. 2.3e-12;					
Matches 246; Conservative		0; Mismatches 230; Indels 0; Gaps 0;					
Qy	33	GCCGCTCCAGCGACGGTGTGAGCATCACTCCGCTCACCTTCTTCGACATGACGTG	92				
Dd	1	GCACCTCCAGCGCGGACGAGGTGTACTTCCTCTTACTTATTTTGACCATGTTTG	60				
Qy	93	GCTGCATTTCACGCCCATGCTTCAGCTTCTCTTACGAATCCCCTGTTCCAAACCGC	152				
Dd	61	GTTAGGGTTTCACCTTATTAGCGGATATTATTCTACAAGCTCTCCATTTCCAAATCCGA	120				
Qy	153	CTTCCTCGAAACCGTGTTCGAAACTCAACAATCCTTATCTTAACCCCTCAACACTTT	212				
Dd	121	TTTCGTTCAAAACATATTTCCTCTTTAAAAATTCACTTTCCCTCACTCAACACTA	180				
Qy	213	CTTCCCCCTTCATGCAATCTAATCTACCCCTATCGCGGAGAAAATGCCGAGTTCCG	272				
Dd	181	TATGCCCTTAGCCGGAACGTTGCTTGCCACTAGATACAAACGGATATCTGAGTTAG	240				
Qy	273	GTATCAGAACGGTGACTCGGTTCTTTTCACGATTATGGAGTCTGTGCGAGATCATCCGA	332				
Dd	241	TTATGTGACAGGAGATCTGTGTCTGTTAAATTTTCTGAGACTGATATGGATTTCAAATTA	300				
Qy	333	TTCCGCTCATAAATACTACTCTTTGCCCCCTAGCGACGATTTAAGAATCTCCAGCTGCC	392				
Dd	301	TCTCATTGGCGACCATCGCGCTAATGCTTAAGGATTTTTATCAGTTTGTCTCAAGTTAGG	360				
Qy	393	CCCGATAGTCAGGAATCTGATCGAAATTCGTTCAAGTTTTTACCGTGCAGTCACTCT	452				
Dd	361	GGAACCTAAGGATGCACCCGGGTCCAATTAGCCCGCTCTTAGCCATTCAAGTGACACT	420				
Qy	453	GTTTCCCGGTCCGCGGTGTCTCGGAATAACACGCACACACCGTTTGGCGATG	508				
Dd	421	TTTTCCGAATCTTGGTGATCCATGTTGTTCACTAACCATCATGTTGTGGTGATG	476				
RESULT 14							
BF096979							
LOCUS		BF096979 482 bp mRNA EST 19-OCT-2000					
DEFINITION		EST400516 tomato nutrient deficient roots Lycopersicon esculentum					
cDNA clone		cDNA clone cLEW18020 5' sequence, mRNA sequence.					
ACCESSION		BF096979					
VERSION		BF096979.1 GI:10902689					

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 16:58:18 ; Search time 8904.87 Seconds
(without alignments)
29.529 Million cell updates/sec

Title: US-08-894-356C-22
Perfect score: 17
Sequence: 1 GAYTYGGNTGGGGNAA 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_ba3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_om:*
- 8: gb_ov:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_pl1:*
- 13: gb_pl2:*
- 14: gb_pl3:*
- 15: gb_pl4:*
- 16: em_ba1:*
- 17: em_ba2:*
- 18: em_fun:*
- 19: em_htgo_hum:*
- 20: em_htgo_inv:*
- 21: em_htgo_rod:*
- 22: em_htg_hum1:*
- 23: em_htg_hum2:*
- 24: em_htg_hum3:*
- 25: em_htg_hum4:*
- 26: em_htg_hum5:*
- 27: em_htg_hum6:*
- 28: em_htg_hum7:*
- 29: em_htg_hum8:*
- 30: em_htg_inv1:*
- 31: em_htg_inv2:*
- 32: em_htg_other:*
- 33: em_htg_rod:*
- 34: em_hum1:*
- 35: em_hum2:*
- 36: em_hum3:*
- 37: em_hum4:*
- 38: em_hum5:*
- 39: em_hum6:*
- 40: em_hum7:*
- 41: em_in:*
- 42: em_om:*
- 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vl:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_vl1:*
- 59: gb_vl2:*
- 60: gb_htg1:*
- 61: gb_htg2:*
- 62: gb_htg3:*
- 63: gb_htg4:*
- 64: gb_htg5:*
- 65: gb_htg6:*
- 66: gb_htg7:*
- 67: gb_htg8:*
- 68: gb_htg9:*
- 69: gb_htg10:*
- 70: gb_htg11:*
- 71: gb_htg12:*
- 72: gb_htg13:*
- 73: gb_htg14:*
- 74: gb_htg15:*
- 75: gb_htg16:*
- 76: gb_htg17:*
- 77: gb_htg18:*
- 78: gb_htg19:*
- 79: gb_htg20:*
- 80: gb_htg21:*
- 81: gb_htg22:*
- 82: gb_htg23:*
- 83: gb_htg24:*
- 84: gb_htg25:*
- 85: gb_pr1:*
- 86: gb_pr2:*
- 87: gb_pr3:*
- 88: gb_pr4:*
- 89: gb_pr5:*
- 90: gb_pr6:*
- 91: gb_pr7:*
- 92: gb_pr8:*
- 93: gb_pr9:*
- 94: gb_rol:*
- 95: gb_rod:*
- 96: gb_in4:*
- 97: gb_pr10:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.2	83.5	850	13	AF227981
2	14.2	83.5	915	53	CNS06W8Q
3	14.2	83.5	1080	9	A16155
4	14.2	83.5	1080	9	A20628
5	14.2	83.5	1080	9	AR003699
6	14.2	83.5	1080	10	I28271
7	14.2	83.5	1096	9	AR110778
8	14.2	83.5	1320	12	AF053307
					AF227981 Euphorbia
					AL418128 T3 end of
					AL6155 PTOM36 10/
					A20628 PTOM36 frag
					AR003699 Sequence
					I28271 Sequence 1
					AR110778 Sequence
					AF053307 Catharant

9	14.2	83.5	1388	13	AF297618	Taxus cus						
10	14.2	83.5	1419	13	AF190130	Taxus cus						
11	14.2	83.5	1436	9	AX025512	Sequence						
12	14.2	83.5	1471	9	AX025518	Sequence						
13	14.2	83.5	1476	12	AB029340	Perilla f						
14	14.2	83.5	1479	10	E12756	Perilla ocl						
15	14.2	83.5	1508	10	E12757	Senecio cru						
16	14.2	83.5	1512	14	CPMWEL2	Z70521 C.melo mrna						
17	14.2	83.5	1526	9	A65937	Sequence 1						
18	14.2	83.5	1542	10	I16758	Sequence 1						
19	14.2	83.5	1563	14	NTHSR201	X95343 N.tabacum m						
20	14.2	83.5	1605	10	E12755	Petunia x						
21	14.2	83.5	1605	12	AB026495	Petunia x						
22	14.2	83.5	1622	10	E12754	Gentiana						
23	14.2	83.5	1622	12	AB026494	Gentiana						
24	14.2	83.5	1679	12	AB010708	Gentiana						
25	14.2	83.5	1703	10	E12753	Gentiana						
26	14.2	83.5	1735	14	DCHCBT3	Z84571 D.caryophyl						
27	14.2	83.5	3706	10	I16766	Sequence 16						
28	14.2	83.5	4746	8	AF228714	Ictalurus						
29	14.2	83.5	8222	3	LLA250129	Lactococc						
30	14.2	83.5	10238	3	U32728	Haemophilus						
31	14.2	83.5	11148	1	AE006520	Streptococ						
32	14.2	83.5	12850	2	AF282249	Lactococc						
33	14.2	83.5	16078	65	AC019969	Drosophil						
34	14.2	83.5	17500	64	AC017179	Drosophil						
35	14.2	83.5	59888	62	AC011588	Homo sapi						
36	14.2	83.5	66237	12	AB016892	Arabidops						
37	14.2	83.5	69348	65	AC020525	Arabidops						
38	14.2	83.5	71494	77	AC090191	Homo sapi						
39	14.2	83.5	74342	12	AB020742	Arabidops						
40	14.2	83.5	78174	4	AC004280	Drosophil						
41	14.2	83.5	79676	12	AB013396	Arabidops						
42	14.2	83.5	82360	13	AP000606	Arabidops						
43	14.2	83.5	83650	12	AB023041	Arabidops						
44	14.2	83.5	87581	14	ATT7H20	Arabidops						
45	14.2	83.5	87841	12	AB005247	Arabidops						
ALIGNMENTS												
RESULT 1												
AF227981	AF227981	850 bp	mrna	PLN	17-FEB-2000							
LOCUS	Euphorbia esula F21J9.20-like protein mRNA, partial cds.											
DEFINITION	Euphorbia esula F21J9.20-like protein mRNA, partial cds.											
ACCESSION	AF227981											
VERSION	AF227981.1 GI:6984225											
KEYWORDS	leafy spurge.											
SOURCE	Euphorbia esula											
ORGANISM	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1; Malpighiales; Euphorbiaceae; Euphorbia.											
REFERENCE	1 (bases 1 to 850)											
AUTHORS	Anderson,J.V. and Horvath,D.P.											
TITLE	Identification of mRNAs expressed in underground adventitious buds of Euphorbia esula (leafy spurge)											
JOURNAL	Unpublished											
AUTHORS	Anderson,J.V. and Horvath,D.P.											
TITLE	Direct Submission											
JOURNAL	Submitted (26-JAN-2000) Plant Science, USDA/ARS, 1605 Albrecht Blvd., Fargo, ND 58105, USA											
FEATURES	Location/Qualifiers											
source	1..850											
	/organism="Euphorbia esula"											
	/db_xref="taxon:3993"											
	/clone="12a"											
	/tissue_type="underground adventitious buds"											
	/dev_stage="3-day induced (defoliated)"											
	<1..660											
CDS	/note="similar to Arabidopsis thaliana F21J9.20 encoded by											
GenBank Accession Number AC000103"												
/codon_start=1												
/evidence=not_experimental												
/product="F21J9.20-like protein"												
/protein_id="AAF34801.1"												
/db_xref="GI:6984226"												
/translation="KRFVFDGSGKIGALKEIAASAIAPKPRVREAVSGLIWKGIIATFK												
KSNPLIRPSVMSVSNLRPTPTPPVPAEAKLSKDIDFEMSGWCRFGLYDADEGNG												
MQDFVNYVKYQGEDGVAICEFGKFAEKALSKIDFEMSGWCRFGLYDADEGNG												
LPTWLISVSTNIRNVCILLDTKDGEGFEAWITLSEDSMWSFESDERVLEFAQVNPVGT												
L"												
3'UTR 661..850												
polyA_site 827												
BASE COUNT 258 a 123 c 227 g 242 t												
ORIGIN												
Query Match 83.5%; Score 14.2; DB 13; Length 850;												
Best Local Similarity 76.5%; Pred. No. 9.2e+02;												
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;												
QY	1	GATTTCGGTGGG	NAA	17								
:												
Db	466	GATTTCGGTGGG	GAA	482								
RESULT 2												
CNS06W8Q/C	915 bp	DNA	STS	11-JAN-2001								
LOCUS	T3 end of clone AX0AA029C10 of library AX0AA from strain CBS 7064											
DEFINITION	of Pichia farinosa, sequence tagged site.											
ACCESSION	AL418128											
VERSION	AL418128.1 GI:12200581											
KEYWORDS	Pichia farinosa.											
SOURCE	Pichia farinosa											
ORGANISM	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Pichia.											
REFERENCE	1 (bases 1 to 915)											
AUTHORS	de Montigny,J., Spehner,C., Souciet,J., Tekala,F., Dujon,B., Wincker,P., Artiguenave,F. and Potier,S.											
TITLE	Genomic Exploration of the Hemiascomycetous Yeasts: 15. Pichia sorbitophila											
JOURNAL	FEBS Lett. 487 (1), 87-90 (2000)											
PUBMED	11152890											
AUTHORS	2 (bases 1 to 915)											
	Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.											
TITLE	Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of yeast species for molecular evolution studies(1)											
JOURNAL	FEBS Lett. 487 (1), 3-12 (2000)											
PUBMED	11152876											
REFERENCE	3 (bases 1 to 915)											
AUTHORS	Genoscope.											
TITLE	Direct Submission											
JOURNAL	Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)											
COMMENT	This STS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.											
FEATURES	Location/Qualifiers											

source		1. .915		/organism="Pichia farinosa"		Query Match		83.5%; Score 14.2; DB 9;		Length 1080;	
		/strain="CBS 7064"				Best Local Similarity		76.5%; Pred. No. 9e+02;			
		/db_xref="taxon:4920"				Matches		13; Conservative		2; Mismatches	
		/clone="AX0A029C10"								0; Gaps	
		/clone_lib="AX0AA"									
		/note="end : T3"									
STS		1. .915									
BASE COUNT		267 a		202 c		146 g		298 t		2 others	
ORIGIN											
<hr/>											
Query Match		83.5%; Score 14.2; DB 53;		Length 915;							
Best Local Similarity		76.5%; Pred. No. 9.1e+02;									
Matches		13; Conservative		2; Mismatches		2; Indels		0; Gaps		0;	
<hr/>											
QY		1 GAYTTYGGNTGGGNAA 17									
Db		149 GACTTTGGCTGGGAAA 133									
<hr/>											
RESULT		3									
LOCUS		A16155		1080 bp		DNA		PAT		03-OCT-1994	
DEFINITION		PTOM36.									
ACCESSION		A16155									
VERSION		A16155.1		GI:640912							
KEYWORDS				synthetic construct.							
SOURCE				synthetic construct.							
ORGANISM				artificial sequence.							
REFERENCE		1		(bases 1 to 1080)							
AUTHORS											
TITLE				TRANSGENIC PLANTS WITH INCREASED SOLIDS CONTENT							
JOURNAL				Patent: WO 9314212-A 1 22-JUL-1993;							
FEATURES				Location/Qualifiers							
source				1. .1080							
				/organism="synthetic construct"							
				/db_xref="taxon:32630"							
BASE COUNT		356 a		141 c		208 g		375 t			
ORIGIN											
<hr/>											
Query Match		83.5%; Score 14.2; DB 9;		Length 1080;							
Best Local Similarity		76.5%; Pred. No. 9e+02;									
Matches		13; Conservative		2; Mismatches		2; Indels		0; Gaps		0;	
<hr/>											
QY		1 GAYTTYGGNTGGGNAA 17									
Db		866 GATTTGGATGGGAAA 882									
<hr/>											
RESULT		4									
LOCUS		A20628		1080 bp		DNA		PAT		12-JUN-1994	
DEFINITION		PTOM36 fragment.									
ACCESSION		A20628									
VERSION		A20628.1		GI:583390							
KEYWORDS				synthetic construct.							
SOURCE				synthetic construct.							
ORGANISM				artificial sequence.							
REFERENCE		1		(bases 1 to 1080)							
AUTHORS											
TITLE				DNA, CONSTRUCTS, CELLS AND PLANTS DERIVED THEREFROM							
JOURNAL				Patent: WO 9105865-A 1 02-MAY-1991;							
FEATURES				Location/Qualifiers							
source				1. .1080							
				/organism="synthetic construct"							
				/db_xref="taxon:32630"							
BASE COUNT		356 a		141 c		208 g		375 t			
ORIGIN											
<hr/>											
Query Match		83.5%; Score 14.2; DB 10;		Length 1080;							
Best Local Similarity		76.5%; Pred. No. 9e+02;									
Matches		13; Conservative		2; Mismatches		2; Indels		0; Gaps		0;	
<hr/>											
QY		1 GAYTTYGGNTGGGNAA 17									
Db		866 GATTTGGATGGGAAA 882									
<hr/>											
RESULT		5									
LOCUS		AR003699		1080 bp		DNA		PAT		04-DEC-1998	
DEFINITION		Sequence 1 from patent US 5744364.									
ACCESSION		AR003699									
VERSION		AR003699.1		GI:3964958							
KEYWORDS				Unknown.							
SOURCE				Unknown.							
ORGANISM				Unclassified.							
REFERENCE		1		(bases 1 to 1080)							
AUTHORS				Bird,C.Roger, Grierson,D., Ray,J.Anthony and Schuch,W.Walter.							
TITLE				PTOM36 constructs and tomato cells transformed therewith							
JOURNAL				Patent: US 5744364-A 1 28-APR-1998;							
FEATURES				Location/Qualifiers							
source				1. .1080							
				/organism="unknown"							
BASE COUNT		356 a		141 c		208 g		375 t			
ORIGIN											
<hr/>											
Query Match		83.5%; Score 14.2; DB 9;		Length 1080;							
Best Local Similarity		76.5%; Pred. No. 9e+02;									
Matches		13; Conservative		2; Mismatches		2; Indels		0; Gaps		0;	
<hr/>											
QY		1 GAYTTYGGNTGGGNAA 17									
Db		866 GATTTGGATGGGAAA 882									
<hr/>											
RESULT		6									
LOCUS		I28271		1080 bp		DNA		PAT		06-FEB-1997	
DEFINITION		Sequence 1 from patent US 5569829.									
ACCESSION		I28271									
VERSION		I28271.1		GI:1819047							
KEYWORDS				Unknown.							
SOURCE				Unknown.							
ORGANISM				Unclassified.							
REFERENCE		1		(bases 1 to 1080)							
AUTHORS				Bird,C.R., Boniwell,J.M., Grierson,D., Ray,J.A. and Schuch,W.W.							
TITLE				Transformed tomato plants							
JOURNAL				Patent: US 5569829-A 1 29-OCT-1996;							
FEATURES				Location/Qualifiers							
source				1. .1080							
				/organism="unknown"							
BASE COUNT		356 a		141 c		208 g		375 t			
ORIGIN											
<hr/>											
Query Match		83.5%; Score 14.2; DB 10;		Length 1080;							
Best Local Similarity		76.5%; Pred. No. 9e+02;									
Matches		13; Conservative		2; Mismatches		2; Indels		0; Gaps		0;	
<hr/>											
QY		1 GAYTTYGGNTGGGNAA 17									
Db		866 GATTTGGATGGGAAA 882									
<hr/>											
RESULT		7									
LOCUS		AR110778		1096 bp		DNA		PAT		14-FEB-2001	

```

DEFINITION Sequence 1 from patent US 6127113.
ACCESSION AR110778
VERSION AR110778.1 GI:12827626
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1096)
AUTHORS Atkinson,R.L. and Dhurandhar,N.V.
TITLE Viral obesity methods and compositions
JOURNAL Patent: US 6127113-A 1 03-OCT-2000;
FEATURES
    source
        location/Qualifiers
            1..1096
                /organism="unknown"
BASE COUNT 353 a 215 c 227 g 301 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 9; Length 1096;
Best Local Similarity 76.5%; Pred. No. 9e+02;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGGNAA 17
||:|||||
Db 1013 GATTTGGATGGGTAA 1029

RESULT 8
AF053307 1320 bp DNA PLN 01-JAN-1999
LOCUS catharanthus roseus deacetylindoline 4-O-acetyltransferase (DAT)
DEFINITION gene, complete cds.
ACCESSION AF053307
VERSION AF053307.1 GI:4091807
KEYWORDS Madagascar periwinkle.
SOURCE Catharanthus roseus
ORGANISM Magnoliophyta; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
I; Gentianales; Apocynaceae; Catharanthus.
REFERENCE 1 (bases 1 to 1320)
AUTHORS St-Pierre,B., Lafamme,P., Alarco,A.M. and De Luca,V.
TITLE The terminal O-acetyltransferase involved in vindoline biosynthesis
defines a new class of proteins responsible for coenzyme
A-dependent acyl transfer
JOURNAL Plant J. 14 (6), 703-713 (1998)
MEDLINE 98346012
REFERENCE 2 (bases 1 to 1320)
AUTHORS St-Pierre,B., Lafamme,P. and De Luca,V.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-1998) Institut de Recherche en Biologie Vegetale,
4101 Sherbrooke East, Montreal, QC H1X 2B2, Canada
FEATURES
    source
        location/Qualifiers
            1..1320
                /organism="Catharanthus roseus"
                /cultivar="Little Delicata"
                /db_xref="taxon:4058"
                <1..>1320
                    /gene="DAT"
                    /product="deacetylindoline 4-O-acetyltransferase"
                    <1..>1320
                        /gene="DAT"
                        <1..>1320
                            /gene="DAT"
                            /EC_number="2.3.1.107"
                            /function="involved in the last step in vindoline
                                biosynthesis"
                            /note="acetylcoenzyme A:deacetylindoline
                                17-O-acetyltransferase"
                                /codon_start=1
                                /product="deacetylindoline 4-O-acetyltransferase"
                                /protein_id="AAC99311.1"
                                /db_xref="GI:4091808"
                                /translation="MESGKISVETETLSKTLIKPSSPTQSLSRYNLSYNDQNIYQTC
                                    VSVGFVENPDGIEISTIRQLQSLSKTLVSYYPAGKVKVKNNDYIHCNDDGIEFVHV
                                    RIRCRNMDILKVELRSYADLVLPKRVTVGSEDTAIVOLSHDFDCGLAVAFGISHKV
                                    ADGGTIASEMKDMAASACYLSSSHHVPPLLVSDSIFPRQDNLICQFPTSKNCVEKT
                                    FIFPPATEKLKSKAVEFIEKTRVETLTAFLSRCATVAGKSAANNKCGSLPPPV
                                    LQAINLRPILLEPQNSVGNLSIYFSRTIKENDYLNEKEYTKLIVINELRKEKQIKRNL
                                    QRYKTKVQAMEEFVKSLEFSDISNFDLIDAYLSDSWCRFPDYDVGFGKPIWVCILF
                                    QPYIKNCVVMMDYFPGDDYIGIEAIVSFQEKMSAFKNEQLQLQFVSN"
BASE COUNT 409 a 244 c 252 g 415 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 12; Length 1320;
Best Local Similarity 76.5%; Pred. No. 8.8e+02;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGGNAA 17
||:|||||
Db 1138 GATTTGGATGGGAAA 1154

RESULT 9
AF297618 1388 bp mRNA PLN 06-DEC-2000
LOCUS Taxus cuspidata 2-debenzoyl-7,13-diacetylbaaccatin III-2-O-benzoyl
DEFINITION transferase mRNA, complete cds.
ACCESSION AF297618
VERSION AF297618.1 GI:11559715
KEYWORDS Taxus cuspidata.
SOURCE Taxus cuspidata
ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Coniferopsida; Coniferales; Taxaceae; Taxus.
REFERENCE 1 (bases 1 to 1388)
AUTHORS Walker,K. and Croteau,R.
TITLE Taxol biosynthesis: Molecular cloning of a benzoyl- CoA:taxane
2alpha -O-benzoyltransferase cDNA from Taxus and functional
expression in Escherichia coli
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (25), 13591-13596 (2000)
MEDLINE 11095755
REFERENCE 2 (bases 1 to 1388)
AUTHORS Walker,K.D. and Croteau,R.B.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-2000) Institute of Biological Chemistry,
Washington State University, PO Box 646340, Pullman, WA 99164-6340,
USA
FEATURES
    source
        location/Qualifiers
            1..1388
                /organism="Taxus cuspidata"
                /db_xref="taxon:99806"
                22..1344
                    /protein_id="AAG38049.1"
                    /db_xref="GI:11559716"
                    /note="DBBT"
                    /codon_start=1
                    /product="2-debenzoyl-7,13-diacetylbaaccatin
                        III-2-O-benzoyl transferase"
                        /translation="MGRFNVDMIERVIVAPCLQSPKNILHLSIDNKRGLTNLSVY
                            NASQSVSADPAKTRIALSKVLVYPPFAGRLRNTENGDLVECTGEGAVEEAMA
                            DNDLSVLQDFNEYPDSPQQLVFNLRDENVIEDLHLLTVQVTRFCGFGVYTRFHHSV
                            SDGKIGICQLKGMGEMARGFKPSLEPIWNREMKVPEDIMYLQDFHDFTHPPLNLEK
                            SIOASNVISFERINYIKRCMMECKEKFSSVVALIWLARTKFRIPDPNVKTIIF
                            PIDMNSFDSPLPKGYGNAIGNACADNVKDLGSLLYALMLIKKSKFALNENKS
                            RILTKESTLDANKKHENVYCCGWRNLGFYADFNGNNAVNSPMQOQREHMLAMQNY
                            FLFLRAKNMIDIGIKILMFMPASVMPAPKIEMEVTINKYVAKICNSKL"
BASE COUNT 446 a 243 c 301 g 398 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 13; Length 1388;
Best Local Similarity 76.5%; Pred. No. 8.8e+02;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY	1	GAYTTYGGNTGGGNAA 17																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
----	---	---------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 5 18-MAR-1997;
SUNTORY LTD
COMMENT OS Senecio cruentus
PN JP 1997070290-A/5
PD 18-MAR-1997
PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO.
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAAKI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,C12N9/10, PC
C12R1:865),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear; Location/Qualifiers
FH Key
FH
FT source 1..1508
FT /organism='Senecio cruentus'
FT /tissue_type='petal'
FT /clone='pCAY48'
FT CDS 1..1367
FT /product='acyltransferase'.
FT Location/Qualifiers
source 1..1508
/organism='unidentified'
/db_xref='taxon:32644'

BASE COUNT 442 a 293 c 296 g 477 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 10; Length 1508;
Best Local Similarity 76.5%; Pred. No. 8.7e+02;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNAA 17
||:|||||
Db 1182 GATTTTGGGTGGGGGAA 1198

Search completed: November 5, 2001, 16:58:20
Job time: 11270 sec

THIS PAGE BLANK (USPTO)


```

XX PA (SUNR ) SUNTORY LTD.
XX PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
XX PI Nakao M, Tanaka Y, Yonekura K;
XX XX WPI; 1996-393401/39.
XX DR DNA coding for aromatic acyl transferase - for transforming plants
XX PT which produce anthocyanin pigments and thus altering colour tone,
XX PT e.g. of flowers
XX XX
XX PS Claim 3; Page 82; 94pp; Japanese.
XX XX
XX CC Vectors containing DNA fragments encoding proteins of plant origin
XX CC with aromatic acyl transferase activity may be used to transform
XX CC plants which produce anthocyanin pigments. The aromatic acyl
XX CC transferase acylates and allowing new colourations to be produced. Six
XX CC tone changes and allowing new colourations to be produced. Six
XX CC specific DNA sequences encoding aromatic acyl transferase from
XX CC different plants are described in AAT37308-T37313. This degenerate
XX CC primer was synthesised base on a peptide fragment (AAW04728)
XX CC isolated from three of the six clones. It was used to identify
XX CC other aromatic acyl transferase encoding clones.
XX XX
XX SQ Sequence 17 BP; 3 A; 0 C; 7 G; 3 T; 4 other;

Query Match 83.5%; Score 14.2; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
Db 1 gaytvygntgggnaa 17

RESULT 2
AAA13989/C
ID AAA13989 standard; DNA; 23 BP.
XX AC AAA13989;
XX XX
XX DT 08-AUG-2000 (first entry)
XX DE
XX KW Geranylgeranyl diphosphate synthase PCR primer SEQ ID NO:6.
XX KW Geranylgeranyl diphosphate synthase; GGPP synthase; yew; cytostatic;
XX KW anticancer; Taxus; diterpene; pacitaxel; identification; plant;
XX KW Taxomyces andreanae; Penicillium raistrickii; microorganism;
XX KW PCR primer; ss.
XX OS Taxus canadensis.
XX XX
XX PN US6043072-A.
XX XX
XX PD 28-MAR-2000.
XX XX
XX PF 05-NOV-1998; 98US-0187050.
XX XX
XX PR 05-NOV-1998; 98US-0187050.
XX XX
XX PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX XX
XX PI Croteau RB, Hefner JL;
XX XX
XX DR WPI; 2000-282526/24.
XX XX
XX PT Nucleic acid encoding geranylgeranyl diphosphate is useful for
XX PT producing pacitaxel and other diterpenes that are useful as anticancer
XX PT drugs
XX XX
XX PS Example 1; Column 39; 57pp; English.

```

```

XX CC The present sequence represents a PCR primer for a geranylgeranyl
XX CC diphosphate (GGPP) synthase protein. GGPP synthase has cytostatic
XX CC activity. A vector encoding GGPP synthase is useful in increasing
XX CC GGPP synthase levels in a host cell preferably faxus (fex) cell and
XX CC thereby facilitates production, isolation and purification of larger
XX CC amounts of GGPP synthase in plants. GGPP synthase is useful in obtaining
XX CC expression or enhanced expression of GGPP and other diterpenes, such as
XX CC pacitaxel, useful as anticancer drugs. Isolated nucleic acids encoding
XX CC GGPP synthase or hybridising with GGPP synthase encoding nucleic acids
XX CC are used for identifying genes encoding GGPP synthase from
XX CC microorganisms such as Taxomyces andreanae and Penicillium raistrickii.
XX XX
XX SQ Sequence 23 BP; 6 A; 9 C; 4 G; 4 T; 0 other;

Query Match 83.5%; Score 14.2; DB 21; Length 23;
Best Local Similarity 76.5%; Pred. No. 51;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
Db 20 GAYTTCGGCTGGGGTAA 4

RESULT 3
AAQ11802
ID AAQ11802 standard; DNA; 1080 BP.
XX AC AAQ11802;
XX XX
XX DT 23-JUL-1991 (first entry)
XX DE Clone pTOM36.
XX XX
XX KW Fruit ripening; pTOM36; ss.
XX OS Lycopersicon esculentum.
XX PN WO9105865-A.
XX PD 02-MAY-1991.
XX PF 17-OCT-1990; 90WO-GB01603.
XX PR 20-OCT-1989; 89GB-0023716.
XX PA (ICIL ) IMPERIAL CHEM INDS PLC.
XX PI Bird CR, Grierson D, Ray JA, Schuch WW;
XX XX
XX DR WPI; 1991-148743/20.
XX XX
XX PT DNA constructs contg. DNA from pTOM36 clone - used to transform
XX PT plants to regulate prodn. of the fruit-ripening pTOM36 enzyme
XX XX
XX PS Disclosure; fig 1; 22pp; English.
XX XX
XX CC This clone is contained in a DNA construct used to transform host
XX CC plant cells for regulating the prodn. of the enzyme encoded by pTOM36.
XX CC Plants such as apple, tomato and mango may be modified to produce e.g.
XX CC sweeter fruit, novel flavour, modified colour or to have improved
XX CC processing characteristics. The DNA construct pref. also comprises
XX CC a constitutive or regulatory promoter, e.g. the CaMV 35S or poly-
XX CC galacturonase gene promoters.
XX XX
XX SQ Sequence 1080 BP; 357 A; 141 C; 207 G; 375 T; 0 other;

Query Match 83.5%; Score 14.2; DB 12; Length 1080;
Best Local Similarity 76.5%; Pred. No. 74;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 GAYTTYGGNTGGGNAA 17
 ||:|||| |

Db 866 gattttggatggggaaa 882

RESULT 4

AAQ46682
 ID AAQ46682 standard; cDNA; 1080 BP.

XX AC AAQ46682;

XX DT 18-JAN-1994 (first entry)

XX DE PTOM36 cDNA clone.

XX KW Transgenic fruit production; prodn.; increased solids content;
 higher reducing sugar content; tomato; tomatoes; plant cultivation;
 improved; paste; soup; ss.

XX OS Synthetic.

XX PN W09314212-A.

XX PD 22-JUL-1993.

XX PF 08-JAN-1993; 93WO-GB000021.

XX PR 10-JAN-1992; 92GB-0000520.

XX PA (ZONE) ZENECA LTD.

XX PI Bird CR, Boniwell JW;

XX DR WPI; 1993-243227/30.

XX PT Transgenic fruit with increased solids content - is obtd. by
 cultivating plants, esp. tomatoes, in which expression of genes
 homologous to PTOM36 is inhibited

XX PS Disclosure; Fig 1; 26pp; English.

XX CC The sequence is that of the cDNA clone pTOM36 which is believed to
 encode a cytoplasmic protein of approximately 52000 daltons
 involved in the ripening of tomatoes. The clone is 1069 bases long
 with an open reading frame of 271 codons. DNA homologous to pTOM36
 may be used in a construct that inhibits expression of genes
 homologous to pTOM36 during ripening. Fruit from plants transformed
 with this construct have increased solids content and a higher
 content of reducing sugars, e.g. glucose and fructose. The fruit
 may be mangoes, peaches, apples, pears, strawberries, bananas,
 melons and esp. tomatoes, which are useful in the prodn. of
 improved tomato paste and soup.

XX SQ Sequence 1080 BP; 356 A; 141 C; 208 G; 375 T; 0 other;

Query Match 83.5%; Score 14.2; DB 14; Length 1080;
 Best Local Similarity 76.5%; Pred. No. 74;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGNTGGGNAA 17
 ||:|||| |

Db 866 gattttggatggggaaa 882

RESULT 5

AAV07891
 ID AAV07891 standard; cDNA; 1096 BP.

XX AC AAV07891;

XX DT 29-JAN-1999 (first entry)

XX

DE Nucleotide sequence of the fiber protein encoded by adenovirus Ad-36p.

XX Adenovirus; Ad-36p fiber protein; viral obesity; vaccine; ss.

XX OS Mastadenovirus.

XX FH Key Location/Qualifiers
 XX FT CDS 1..1096
 XX FT /*tag= a
 XX FT /product= "Ad-36p fiber protein"

XX PN W09844946-A1.

XX PD 15-OCT-1998.

XX PF 06-APR-1998; 98WO-US06730.

XX PR 04-APR-1997; 97US-0042942.

XX PA (OBET-) OBETECH LLC.

XX PI Atkinson RL, Dhurandhar NV;

XX DR WPI; 1998-568305/48.

XX PT Determining if obesity in a person is caused by Ad-36 virus - and
 providing the basis for treatment or prevention of obesity-causing,
 cholesterol reducing adenovirus, using the purified variant, Ad-36p

XX PS Claim 1; Pages 18-19; 24pp; English.

XX CC This is the nucleotide sequence encoding the adenovirus Ad-36p fiber
 protein used in the method of the invention to determine if a person
 is suffering viral obesity. The method is used to determine whether
 obesity in a person has a viral basis. Ad-36p can be used as a basis
 of a vaccine to prevent viral-based obesity.

XX SQ Sequence 1096 BP; 353 A; 215 C; 227 G; 301 T; 0 other;

Query Match 83.5%; Score 14.2; DB 19; Length 1096;
 Best Local Similarity 76.5%; Pred. No. 74;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGNTGGGNAA 17

Db 1013 gattttggatggggtaa 1029

RESULT 6

AAC49764
 ID AAC49764 standard; DNA; 1427 BP.

XX AC AAC49764;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 62347.

XX KW Hybridisation assay; genetic mapping; gene expression control;
 protein identification; signal transduction pathway;
 metabolic pathway; promoter; termination sequence; ss.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143342.
 PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 23-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 22-SEP-1999; 99US-0154779.
 PR 23-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 83.5%; Score 14.2; DB 21; Length 1427;
Best Local Similarity 76.5%; Pred. No. 76;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGNA 17
||:|||||
Db 1161 gatttcggttggttaa 1177

RESULT 7
AAC33965
ID AAC33965 standard; DNA; 1429 bp.
XX AC AAC33965;
XX DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 4961.
XX DE
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD
XX 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140833.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

```
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147937.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 28-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 03-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.

PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 83.5%; Score 14.2; DB 21; Length 1429;
Best Local Similarity 76.5%; Pred. No. 76;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGTGGGNA 17
||:|:|:|:|:|:|:|
Db 1163 gattcggttggggttaa 1179

RESULT 8
AAC64784
ID AAC64784 standard; cDNA; 1436 BP.
XX
AC AAC64784;
XX
DT 28-FEB-2001 (first entry)
XX
DE Lemon alcohol acyl transferase encoding cDNA SEQ ID NO:20A.
XX
KW Strawberry; fruit flavour; biosynthetic pathway; aliphatic; thiolase;
KW aromatic ester; alcohol acyl transferase; alcohol dehydrogenase;
KW pyruvate decarboxylase; aminotransferase; esterase; alcohol; aldehyde;
KW alpha-keto acid; amino acid; fatty acid; acyl-CoA; processed food;
KW food additive; flavour; syrup; ice-cream; frozen dessert; yoghurt;
KW confectionery; flavouring; oral medication; vitamin; aroma; beverage;
KW alcohol; scent; fragrance; perfume; cosmetic; suspension aid;
KW aluminium salt; anti-perspirant; pharmaceutical; cleaning product;
KW insect pheromone; dye carrier; solvent; insect repellent; miticide;
KW scabicide; plasticiser; deodorant; ss.
XX
OS Citrus limon.
XX
PN WO200032789-A1.
XX
PD 08-JUN-2000.
XX
PF 02-DEC-1999; 99WO-NL00737.
XX
XX 02-DEC-1998; 98EP-0204018.
PR 12-MAR-1999; 99EP-0200739.
XX
PA (CPRO-) CPRO-DLO CENT PLANTENVEREDELINGS REPROD.
XX
XX Aharoni A, Luecker J, Verhoeven HA, Van Tunen AJ, O'Connell AP;
PI WPI; 2000-412335/35.
XX P-PSDB; AAB36458.
DR
```


XX A new DNA sequence encoding a polypeptide with alcohol acyl transferase
PT activity for producing and regulating aromatic and/or aliphatic ester
PT formation in microorganisms, plant cells or plants -
XX Example 5; Page 114-115; 163pp; English.
XX
CC The present invention describes nucleotide sequences with thiolase,
CC alcohol acyl transferase, alcohol dehydrogenase, pyruvate decarboxylase,
CC aminotransferase and esterase activities, which are involved in the
CC biosynthetic pathway for aliphatic and/or aromatic ester production in
CC fruit. The nucleotide sequences can be inserted into the genome of a
CC fruit-producing plant to regulate aliphatic and/or aromatic ester
CC formation. Aromatic and/or aliphatic esters in microorganisms, plant
CC cells or plants are produced by inserting thiolase, alcohol acyl
CC transferase, alcohol dehydrogenase, pyruvate decarboxylase,
CC aminotransferase and esterase nucleotide sequences into the genome and
CC feeding the microorganism or plant with alcohol, aldehydes, alpha-keto
CC acids, or amino acids and fatty acids, and acyl-CoA. The nucleotides and
CC their proteins can be used in the processed food industry as food
CC additives to enhance the flavour of syrups, ice-creams, frozen desserts,
CC yoghurts and confectionery. They are used: as flavouring agents for oral
CC medications and vitamins; provide flavour and aroma in beverages,
CC including alcohol; enhance or reduce fruit flavour, aroma, fragrance or
CC scent; enhance the flavour or aroma of natural, synthetic or artificial
CC products; for the production of novel combinations of artificial flavour
CC substances; as antibacterial or anti-fungal agents; as fragrance or
CC perfumes in cosmetics, creams, sun-protectant products, hair
CC conditioners, lengthening agents and fixatives in perfumes, suspension
CC aids for aluminum salts in anti-perspirant pharmaceuticals, cleaning
CC products, personal care products and animal care products; as
CC disinfectant additives; as degreasing solvents for electronics; as
CC insect pheromones; and as dye carriers, solvents, insect repellents,
CC miticides, scabicides, plasticisers and deodorants. The present sequence
CC encodes lemon alcohol acyl transferase, from the present invention.
XX
SQ Sequence 1436 BP; 410 A; 280 C; 320 G; 426 T; 0 other;

Query Match 83.5%; Score 14.2; DB 21; Length 1436;
Best Local Similarity 76.5%; Pred. No. 76;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGGNAA 17
DB 1126 gatttcgggtgggggaa 1142
||:|:|:|:|:|:|:|:|:|

RESULT 9
AAC64787
ID AAC64787 standard; cDNA; 1470 BP.
XX
AC AAC64787;
XX
XX 28-FEB-2001 (first entry)
XX
XX Honey dew melon alcohol acyl transferase encoding cDNA SEQ ID NO:23A.
XX
KW Strawberry; fruit flavour; biosynthetic pathway; aliphatic; thiolase;
KW aromatic ester; alcohol acyl transferase; alcohol dehydrogenase;
KW pyruvate decarboxylase; aminotransferase; esterase; alcohol; aldehyde;
KW alpha-keto acid; amino acid; fatty acid; acyl-CoA; processed food;
KW food additive; flavour; syrup; ice-cream; frozen dessert; yoghurt;
KW confectionery; flavouring; oral medication; vitamin; aroma; beverage;
KW alcohol; scent; fragrance; perfume; cosmetic; suspension aid;
KW aluminum salt; anti-perspirant; pharmaceutical; cleaning product;
KW insect pheromone; dye carrier; solvent; insect repellent; miticide;
KW scabicide; plasticiser; deodorant; ss.
XX
XX Cucumis sp.
OS
XX WO200032789-A1.
PN
XX

PD 08-JUN-2000.
XX
XX 02-DEC-1999; 99WO-NL00737.
XX
PR 02-DEC-1998; 98EP-0204018.
PR 12-MAR-1999; 99EP-0200739.
XX
XX (CPRO-) CPRO-DLO CENT PLANTENVERDEDELINGS REPROD.
PA Aharoni A, Luecker J, Verhoeven HA, Van Tunen AJ, O'Connell AP;
PI WPI; 2000-412335/35.
DR P-PSDB; AAB36461.
XX
XX A new DNA sequence encoding a polypeptide with alcohol acyl transferase
PT activity for producing and regulating aromatic and/or aliphatic ester
PT formation in microorganisms, plant cells or plants -
PS Example 5; Page 121-122; 163pp; English.
XX
CC The present invention describes nucleotide sequences with thiolase,
CC alcohol acyl transferase, alcohol dehydrogenase, pyruvate decarboxylase,
CC aminotransferase and esterase activities, which are involved in the
CC biosynthetic pathway for aliphatic and/or aromatic ester production in
CC fruit. The nucleotide sequences can be inserted into the genome of a
CC fruit-producing plant to regulate aliphatic and/or aromatic ester
CC formation. Aromatic and/or aliphatic esters in microorganisms, plant
CC cells or plants are produced by inserting thiolase, alcohol acyl
CC transferase, alcohol dehydrogenase, pyruvate decarboxylase,
CC aminotransferase and esterase nucleotide sequences into the genome and
CC feeding the microorganism or plant with alcohol, aldehydes, alpha-keto
CC acids, or amino acids and fatty acids, and acyl-CoA. The nucleotides and
CC their proteins can be used in the processed food industry as food
CC additives to enhance the flavour of syrups, ice-creams, frozen desserts,
CC yoghurts and confectionery. They are used: as flavouring agents for oral
CC medications and vitamins; provide flavour and aroma in beverages,
CC including alcohol; enhance or reduce fruit flavour, aroma, fragrance or
CC scent; enhance the flavour or aroma of natural, synthetic or artificial
CC products; for the production of novel combinations of artificial flavour
CC substances; as antibacterial or anti-fungal agents; as fragrance or
CC perfumes in cosmetics, creams, sun-protectant products, hair
CC conditioners, lengthening agents and fixatives in perfumes, suspension
CC aids for aluminum salts in anti-perspirant pharmaceuticals, cleaning
CC products, personal care products and animal care products; as
CC disinfectant additives; as degreasing solvents for electronics; as
CC insect pheromones; and as dye carriers, solvents, insect repellents,
CC miticides, scabicides, plasticisers and deodorants. The present sequence
CC encodes honey dew melon alcohol acyl transferase, from the present
XX invention.
SQ Sequence 1470 BP; 415 A; 301 C; 334 G; 420 T; 0 other;

Query Match 83.5%; Score 14.2; DB 21; Length 1470;
Best Local Similarity 76.5%; Pred. No. 77;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGGNAA 17
DB 1122 gactttggatggggaaa 1138
||:|:|:|:|:|:|:|:|:|

RESULT 10
AAC39128
ID AAC39128 standard; DNA; 1477 BP.
XX
XX AAC39128;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 23476.
XX
XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.

[illegible]

PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 83.5%; Score 14.2; DB 21; Length 1477;
 Best Local Similarity 76.5%; Pred. No. 77;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNAA 17
 ||:||||| ||||| ||
 Db 1186 gatttcggttggtggttaa 1202

RESULT 11
 AAT37311
 ID AAT37311 standard; cDNA to mRNA; 1479 BP.
 XX AC AAT37311;
 XX DT 06-FEB-1997 (first entry)
 XX DE Aromatic acyl transferase coding sequence.

XX Aromatic acyl transferase; transformation; anthocyanin pigment;
 KW plants; acylation; colour; tone; colouration; colour change;
 KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;
 KW Scenecio cruentus; Lavandula angustifolia; ds.
 OS Perilla ocimoides (Clone pSAT208).
 XX PN
 XX PD Key Location/Qualifiers

FT CDS 3..1343
 FT /*tag= a
 FT /product= Aromatic acyl transferase.
 XX PN WO9625500-A1.
 XX PD 22-AUG-1996.
 XX PF 16-FEB-1996; 96WO-JP00348.
 XX PR 30-JAN-1996; 96JP-0046534.
 XX PR 17-FEB-1995; 95JP-0067159.
 XX PR 29-JUN-1995; 95JP-0196915.
 XX (SUNR) SUNTORY LTD.
 XX PA Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
 PI Nakao M, Tanaka Y, Yonekura K;
 XX WIPI: 1996-393401/39.
 DR P-PSDB; AAW04725.
 XX DNA coding for aromatic acyl transferase - for transforming plants
 PT which produce anthocyanin pigments and thus altering colour tone,
 PT e.g. of flowers
 XX Claim 4; Page 65-69; 94pp; Japanese.
 XX Vectors containing DNA fragments encoding proteins of plant origin
 CC with aromatic acyl transferase activity may be used to transform
 CC plants which produce anthocyanin pigments. The aromatic acyl
 CC transferase acylates the pigments in the flower resulting in colour
 CC tone changes and allowing new colourations to be produced. Six
 CC specific DNA sequences encoding aromatic acyl transferase from
 CC different plants are described in AAT37308-T37313.
 XX SQ Sequence 1479 BP; 420 A; 316 C; 331 G; 411 T; 1 other;

Query Match 83.5%; Score 14.2; DB 17; Length 1479;
 Best Local Similarity 76.5%; Pred. No. 77;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNAA 17
 ||:||||| ||||| ||
 Db 1167 gatttcggttggtggttaa 1183

RESULT 12
 AAT37312
 ID AAT37312 standard; cDNA to mRNA; 1508 BP.
 XX AC AAT37312;
 XX DT 06-FEB-1997 (first entry);
 XX DE Aromatic acyl transferase coding sequence.

XX Aromatic acyl transferase; transformation; anthocyanin pigment;
 KW plants; acylation; colour; tone; colouration; colour change;
 KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;
 KW Scenecio cruentus; Lavandula angustifolia; ds.
 XX OS Senecio cruentus (Clone pCAT8).
 XX FH Key Location/Qualifiers
 FT CDS 3..1367
 FT /*tag= a
 FT /product= Aromatic acyl transferase.
 XX PN WO9625500-A1.
 XX PD 22-AUG-1996.

```
XX PF 16-FEB-1996; 96WO-JP00348.
XX PR 30-JAN-1996; 96JP-0046534.
XX PR 17-FEB-1995; 95JP-0067159.
XX PR 29-JUN-1995; 95JP-0196915.
XX PA (SUNR ) SUNTORY LTD.
XX PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
XX PI Nakao M, Tanaka Y, Yonekura K;
XX DR WPI; 1996-393401/39.
XX PS P-PSDB; AAW04726.
XX PT DNA coding for aromatic acyl transferase - for transforming plants
XX PT which produce anthocyanin pigments and thus altering colour tone,
XX PT e.g. of flowers
XX PS Claim 4; Page 69-72; 94pp; Japanese.
XX CC Vectors containing DNA fragments encoding proteins of plant origin
XX CC with aromatic acyl transferase activity may be used to transform
XX CC plants which produce anthocyanin pigments. The aromatic acyl
XX CC transferase acylates the pigments in the flower resulting in colour
XX CC tone changes and allowing new colourations to be produced. Six
XX CC specific DNA sequences encoding aromatic acyl transferase from
XX CC different plants are described in AAT37308-T37313.
XX SQ Sequence 1508 BP; 442 A; 294 C; 295 G; 477 T; 0 other;

Query Match 83.5%; Score 14.2; DB 17; Length 1508;
Best Local Similarity 76.5%; Pred. No. 77;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGTGGGNAA 17
Db 1182 gatttgggtggggaa 1198
||:|:|:|:|:|:|

RESULT 13
AAT37313
ID AAT37313 standard; cDNA to mRNA; 1518 BP.
AC AAT37313;
XX DT 06-FEB-1997 (first entry)
XX DE Aromatic acyl transferase coding sequence.
XX KW Aromatic acyl transferase; transformation; anthocyanin pigment;
XX KW plants; acylation; colour; tone; colouration; colour change;
XX KW Gentiana triflora; Petunia hybrida; Perilla occimoides;
XX KW Scenecio cruentus; Lavandula angustifolia; ds.
XX OS Lavandula angustifolia (Clone pLAT21).
XX FH Key Location/Qualifiers
XX FT CDS 1..1355
XX FT /*tag= a
XX FT /product= Aromatic acyl transferase.
XX PN WO9625500-A1.
XX PD 22-AUG-1996.
XX PF 16-FEB-1996; 96WO-JP00348.
XX PR 30-JAN-1996; 96JP-0046534.
XX PR 17-FEB-1995; 95JP-0067159.
XX PR 29-JUN-1995; 95JP-0196915.
XX
```

```
PA (SUNR ) SUNTORY LTD.
XX PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
XX PI Nakao M, Tanaka Y, Yonekura K;
XX DR WPI; 1996-393401/39.
XX PS DNA coding for aromatic acyl transferase - for transforming plants
XX PT which produce anthocyanin pigments and thus altering colour tone,
XX PT e.g. of flowers
XX PS Claim 4; Page 73-76; 94pp; Japanese.
XX CC Vectors containing DNA fragments encoding proteins of plant origin
XX CC with aromatic acyl transferase activity may be used to transform
XX CC plants which produce anthocyanin pigments. The aromatic acyl
XX CC transferase acylates the pigments in the flower resulting in colour
XX CC tone changes and allowing new colourations to be produced. Six
XX CC specific DNA sequences encoding aromatic acyl transferase from
XX CC different plants are described in AAT37308-T37313. NOTE: This
XX CC sequence is supposed to cross reference with the protein described
XX CC in AAW04727, however there are so many discrepancies between the
XX CC polypeptide decoded from this sequence and the polypeptide given in
XX CC the specification and described in AAW04727 that the indexer decided
XX CC not to cross reference the two.
XX SQ Sequence 1518 BP; 384 A; 340 C; 366 G; 428 T; 0 other;

Query Match 83.5%; Score 14.2; DB 17; Length 1518;
Best Local Similarity 76.5%; Pred. No. 77;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGTGGGNAA 17
Db 1176 gatttgggtgggtaa 1192
||:|:|:|:|:|:|

RESULT 14
AAT89415
ID AAT89415 standard; cDNA; 1526 BP.
XX AC AAT89415;
XX DT 15-APR-1998 (first entry)
XX DE Melon ripening-related cDNA clone MEL2.
XX KW Melon cultivar Canteloupe charentais; fruit ripening control;
XX KW ethylene; MEL2 promoter; ss.
XX OS Cucumis melo.
XX FH Key Location/Qualifiers
XX FT 3'UTR 1385..1526
XX FT /*tag= a
XX FT /note= "contains a putative polyadenylation signal"
XX PN WO9737023-A1.
XX PD 09-OCT-1997.
XX PF 24-MAR-1997; 97WO-GB00824.
XX PR 02-APR-1996; 96GB-0006906.
XX PA (ZENE ) ZENECA LTD.
XX PI Aggelis A, Grellerson D, John I, Karvouni Z;
XX PR WPI; 1997-503108/46.
XX PR cDNA related with fruit ripening - can be used to genetically modify
XX
```

```

PT especially melons to control ripening
XX
PS Claim 1: Pages 18-19; 38pp; English.
XX
CC This cDNA sequence is from the MEL2 clone which produces a
CC novel ripening-related product from Cucumis melo. It is one of two
CC specified sequences (MEL2) 1526 bp or (MEL7) 686 bp which were
CC isolated from a melon ripe fruit cDNA library. MEL2 is not a
CC full-length clone; it lacks the initiation codon for the amino terminus.
CC The cDNA allows control of the ripening of fruit, especially melons.
XX
XX Sequence 1526 BP; 439 A; 297 C; 331 G; 448 T; 2 U; 9 other;
SQ

```

Query Match 83.5%; Score 14.2; DB 18; Length 1526;
Best Local Similarity 76.5%; Pred. No. 77;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 GAYTTYGGNTGGGNAA 17
   ||:|||| ||||| ||
Db 1136 gactttggtgggaaa 1152

```

RESULT 15
AAQ54685
ID AAQ54685 standard; cDNA; 1542 BP.
XX
AC AAQ54685;
XX
DT 06-JUL-1994 (first entry)
XX
DE Anther specific cDNA clone ant32.
XX
KW Transgenic plants; male sterility; pollen; sterile;
KW self-pollination; seed; hybrid; toxin-A; peIE; T-urf13; Gin; iaal;
KW CytA; toxin; Nicotiana tabacum; Diptheria; Erwinia chrysanthemi;
KW phage Mu; Psuedomonas syringae; Bacillus thuringiensis; anther;
KW breeding; ss.
XX
OS Nicotiana tabacum.
XX
FH Key Location/Qualifiers
FT CDS 66..1412
FT FT /*tag= a
XX
PN EP578611-A.
XX
PD 12-JAN-1994.
XX
PF 24-JUN-1993; 93EP-0810455.
XX
PR 02-JUL-1992; 92US-0908242.
XX
PA (CIBA) CIBA GEIGY AG.
XX
PI Crossland LD, Tuttle AB;
XX
PI WPI: 1994-010428/02.
DR P-PSDB; AAR47475.
XX
PT Anther-specific cDNA, genomic and recombinant DNA - produce
PT transgenic male-sterile plants, which prevents self-pollination,
PT in hybrid seed prodn.
XX
PS Claim 1; Page 21-24; 75pp; English.
XX
CC Anther specific cDNA or genomic sequences can be used to identify
CC and isolate anther specific promoters. The anther specific promoter
CC can then be cloned into a recombinant construct and used to express
CC heterologous genes. Preferred heterologous genes include Diptheria
CC toxin A-chain gene; pectate lyase gene peIE from Erwinia
CC chrysanthemi; T-urf13 from cms-T maize mitochondrial genomes; the
CC Gin recombinase gene from phage Mu; the indole acetic acid-lysine

```

CC synthetase gene from Pseudomonas syringae and the CytA toxin gene
CC from Bacillus thuringiensis israeliensis. All of these genes when
CC expressed in anther tissue will result in the inability of the
CC plant to produce viable pollen. Transformation of plants with such
CC a recombinant construct can produce transgenic, male sterile plants.
CC Male sterility is important in the production of hybrid seeds as it
CC prevents self pollination which hinders breeding and hybrid seed
CC production.
XX
XX Sequence 1542 BP; 473 A; 293 C; 332 G; 444 T; 0 other;
SQ

```

Query Match 83.5%; Score 14.2; DB 15; Length 1542;
Best Local Similarity 76.5%; Pred. No. 77;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 GAYTTYGGNTGGGNAA 17
   ||:|||| ||||| ||
Db 1239 gattttgggtgggaaa 1255

```

Search completed: November 5, 2001, 18:12:09
Job time: 15464 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:05:10 ; Search time 168.74 Seconds
(without alignments)
19.072 Million cell updates/sec

Title: US-08-894-356C-22

Perfect score: 17

Sequence: 1 GAYTTYGGNTGGGNA 17

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	14.2	83.5	23	US-09-187-050-6	Sequence 6, Appli
2	14.2	83.5	1080	US-07-598-873-1	Sequence 1, Appli
3	14.2	83.5	1080	US-08-073-425-1	Sequence 1, Appli
4	14.2	83.5	1080	US-08-396-531-1	Sequence 1, Appli
5	14.2	83.5	1096	US-09-056-153-1	Sequence 1, Appli
6	14.2	83.5	1512	US-09-142-514-1	Sequence 1, Appli
7	14.2	83.5	1542	US-08-207-904-1	Sequence 1, Appli
8	14.2	83.5	3706	US-08-207-904-16	Sequence 16, Appli
9	13	76.5	4089	US-07-908-245-1	Sequence 1, Appli
C 10	13	76.5	4097	US-09-123-708-5	Sequence 5, Appli
C 11	13	76.5	4097	US-09-123-624-5	Sequence 5, Appli
12	12.6	74.1	508	US-08-822-028-74	Sequence 74, Appli
13	12.6	74.1	508	US-08-478-285-74	Sequence 74, Appli
14	12.6	74.1	912	US-08-593-171A-24	Sequence 24, Appli
15	12.6	74.1	912	US-08-646-590B-24	Sequence 24, Appli
16	12.6	74.1	912	US-09-069-226-24	Sequence 24, Appli
17	12.6	74.1	912	US-09-413-184-24	Sequence 24, Appli
18	12.6	74.1	930	US-08-096-182A-1	Sequence 1, Appli
19	12.6	74.1	930	US-08-877-109-1	Sequence 1, Appli
20	12.6	74.1	930	US-08-798-760-1	Sequence 1, Appli
21	12.6	74.1	930	PCT-US94-08327-1	Sequence 1, Appli
22	12.6	74.1	2344	US-07-695-472B-35	Sequence 35, Appli
23	12.6	74.1	2347	US-08-250-740-32	Sequence 32, Appli
24	12.6	74.1	2347	US-07-695-472B-1	Sequence 1, Appli
25	12.6	74.1	2347	US-07-695-472B-34	Sequence 34, Appli
26	12.6	74.1	2347	US-07-695-472B-36	Sequence 36, Appli
27	12.6	74.1	2560	US-07-916-098A-44	Sequence 44, Appli

28	12.6	74.1	3167	2	US-08-472-534-1	Sequence 1, Appli
C 29	12.6	74.1	3393	1	US-08-295-502-1	Sequence 1, Appli
C 30	12.6	74.1	3393	5	PCT-US95-10691-1	Sequence 1, Appli
C 31	12.6	74.1	4543	2	US-08-519-547A-5	Sequence 5, Appli
32	12.6	74.1	4741	1	US-07-695-472B-4	Sequence 4, Appli
C 33	12.6	74.1	4742	1	US-08-250-740-35	Sequence 35, Appli
C 34	12.2	71.8	531	3	US-08-784-582-59	Sequence 59, Appli
C 35	12.2	71.8	579	1	US-07-749-446-1	Sequence 1, Appli
C 36	12.2	71.8	603	4	US-08-149-101A-5	Sequence 5, Appli
C 37	12.2	71.8	603	5	PCT-US94-12873-5	Sequence 5, Appli
38	12.2	71.8	626	3	US-08-906-769-150	Sequence 150, App
39	12.2	71.8	626	3	US-08-906-616-150	Sequence 150, App
40	12.2	71.8	626	3	US-08-639-075A-150	Sequence 150, App
41	12.2	71.8	626	4	US-09-012-431-150	Sequence 150, App
42	12.2	71.8	626	4	US-09-012-692-150	Sequence 150, App
43	12.2	71.8	626	4	US-08-906-613-150	Sequence 150, App
C 44	12.2	71.8	679	4	US-09-286-132-4	Sequence 4, Appli
C 45	12.2	71.8	782	1	US-07-865-878A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-187-050-6/c
; Sequence 6, Application US/09187050B
; Patent No. 6043072
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Hefner, Jerry
; TITLE OF INVENTION: Nucleic Acids Encoding Taxus Geranylgeranyl Diphosphate
; FILE REFERENCE: WSUR12423
; CURRENT APPLICATION NUMBER: US/09/187,050B
; CURRENT FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 23
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: non-degenerate
; OTHER INFORMATION: reverse PCR primer
; FEATURE:
; NAME/KEY: misc_difference
; LOCATION: (1)..(23)
; OTHER INFORMATION: No. 6043072-degenerate PCR primer
US-09-187-050-6

Query Match 83.5%; Score 14.2; DB 3; Length 23;
Best Local Similarity 76.5%; Pred. No. 8.4;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|||||
DB 20 GATTTGGCTGGGTAA 4

RESULT 2
US-07-598-873-1
; Sequence 1, Application US/07598873
; Patent No. 5254800
; GENERAL INFORMATION:
; APPLICANT: BIRD, COLIN R
; APPLICANT: GRIERSON, DONALD
; APPLICANT: RAY, JOHN A
; APPLICANT: SCHUCH, WOLFGANG W
; TITLE OF INVENTION: DNA, CONSTRUCTS, CELLS AND PLANTS
; TITLE OF INVENTION: DERIVED THEREFROM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

```
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: Eleventh Floor, 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/598,873
; FILING DATE: 19901019
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1080 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Lycopersicon esculentum
; STRAIN: Ailsa Craig
; DEVELOPMENTAL STAGE: Ripening
; US-07-598-873-1

Query Match      83.5%; Score 14.2; DB 1; Length 1080;
Best Local Similarity 76.5%; Pred. No. 14;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GAYTTYGGTGGGNA 17
||:|||||
Db 866 GAYTTGGATGGGAAA 882

RESULT 3
US-08-073-425-1
; Sequence 1, Application US/08073425
; Patent No. 5569829
; GENERAL INFORMATION:
; APPLICANT: BIRD, COLIN R
; APPLICANT: BONIWELL, JEREMY M.
; APPLICANT: GRIERSON, DONALD
; APPLICANT: RAY, JOHN A
; APPLICANT: SCHUCH, WOLFGANG W
; TITLE OF INVENTION: PLANTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 New York Avenue, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/425,425
; FILING DATE: 09-JUN-1993
; CLASSIFICATION: 800

; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: Eleventh Floor, 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/598,873
; FILING DATE: 19901019
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1080 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Lycopersicon esculentum
; STRAIN: Ailsa Craig
; DEVELOPMENTAL STAGE: Ripening
; US-07-598-873-1

Query Match      83.5%; Score 14.2; DB 1; Length 1080;
Best Local Similarity 76.5%; Pred. No. 14;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GAYTTYGGTGGGNA 17
||:|||||
Db 866 GAYTTGGATGGGAAA 882

RESULT 4
US-08-396-531-1
; Sequence 1, Application US/08396531
; Patent No. 5744364
; GENERAL INFORMATION:
; APPLICANT: BIRD, COLIN R
; APPLICANT: GRIERSON, DONALD
; APPLICANT: RAY, JOHN A
; APPLICANT: SCHUCH, WOLFGANG W
; TITLE OF INVENTION: DNA, CONSTRUCTS, CELLS AND PLANTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: Ninth Floor, 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/396,531
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,037
; FILING DATE: 16-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1080 base pairs
; TYPE: nucleic acid
```


STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Lycopersicon esculentum
STRAIN: Ailisa Craig
DEVELOPMENTAL STAGE: Ripening
US-08-396-531-1

Query Match 83.5%; Score 14.2; DB 1; Length 1080;
Best Local Similarity 76.5%; Pred. No. 14;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|:|:|:|:|:|:|:|:|

Db 866 GATTTGGATGGGAAA 882

RESULT 5

US-09-056-153-1
Sequence 1, Application US/09056153
Patent No. 6127113
GENERAL INFORMATION:
APPLICANT: Atkinson, Richard L.
APPLICANT: Dhurandhar, Nikhil V.
TITLE OF INVENTION: Viral Obesity Methods and Compositions
FILE REFERENCE: 710395-90010
CURRENT APPLICATION NUMBER: US/09/056,153
CURRENT FILING DATE: 1998-04-06
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1096
TYPE: DNA
ORGANISM: Adenovirus type 36P
US-09-056-153-1

Query Match 83.5%; Score 14.2; DB 3; Length 1096;
Best Local Similarity 76.5%; Pred. No. 14;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|:|:|:|:|:|:|:|:|

Db 1013 gactttggatggggtaa 1029

RESULT 6

US-09-142-514-1
Sequence 1, Application US/09142514A
Patent No. 6107548
GENERAL INFORMATION:
APPLICANT: Aggelis, Alexandros
APPLICANT: Grierson, Donald
APPLICANT: John, Isaac
APPLICANT: Karvouni, Zoi
TITLE OF INVENTION: Fruit Ripening
FILE REFERENCE: SEES0111/UST
CURRENT APPLICATION NUMBER: US/09/142,514A
CURRENT FILING DATE: 1998-09-09
EARLIER APPLICATION NUMBER: PCT/GB97/00824
EARLIER FILING DATE: 1997-03-24
EARLIER APPLICATION NUMBER: GB 9606906.7
EARLIER FILING DATE: 1996-04-02
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1512
TYPE: DNA
ORGANISM: Cucumis melo
US-09-142-514-1

Query Match 83.5%; Score 14.2; DB 3; Length 1512;
Best Local Similarity 76.5%; Pred. No. 15;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|:|:|:|:|:|:|:|:|

Db 1122 gactttggatggggaaa 1138

RESULT 7

US-08-207-904-1
Sequence 1, Application US/08207904
Patent No. 5477002
GENERAL INFORMATION:
APPLICANT: Tuttle, AnnMarie
APPLICANT: Crossland, Lyle D.
TITLE OF INVENTION: Anther-Specific cDNA Sequences, Genomic
TITLE OF INVENTION: DNA Sequences and Recombinant DNA Sequences
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/207,904
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/908,242
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: CGC 1624
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8615
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1542 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Nicotiana tabacum
INDIVIDUAL ISOLATE: Ant32
FEATURE:
NAME/KEY: CDS
LOCATION: 66..1412
US-08-207-904-1

Query Match 83.5%; Score 14.2; DB 1; Length 1542;
Best Local Similarity 76.5%; Pred. No. 15;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|:|:|:|:|:|:|:|:|

Db 1239 GATTTGGTGGGAAA 1255

RESULT 8
US-08-207-904-16
; Sequence 16, Application US/08207904
; Patent No. 5477002
; GENERAL INFORMATION:
; APPLICANT: Tuttle, AnnMarie
; APPLICANT: Crossland, Lyle D.
; TITLE OF INVENTION: Anther-Specific cDNA Sequences, Genomic
; TITLE OF INVENTION: DNA Sequences and Recombinant DNA Sequences
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/207,904
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/908,242
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: CGC 1624
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8615
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3706 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Nicotiana tabacum
; INDIVIDUAL ISOLATE: Anc32 genomic clone
; IMMEDIATE SOURCE:
; CLONE: pCIB950
; FEATURE:
; NAME/KEY: TATA_signal
; LOCATION: 1971..1975
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2076..3422
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2009
; OTHER INFORMATION: /note= "Putative transcription
; OTHER INFORMATION: start site"
US-08-207-904-16

Query Match 83.5%; Score 14.2; DB 1; Length 3706;
Best Local Similarity 76.5%; Pred. No. 17;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGNA 17
|||:|||||
Db 3249 GATTTGGTGGGAAA 3265

RESULT 9
US-07-908-245-1/c
; Sequence 1, Application US/07908245
; Patent No. 5498539
; GENERAL INFORMATION:
; APPLICANT: Harrison, David G.
; APPLICANT: Alexander, R. Wayne
; APPLICANT: Murphy, T. J.
; APPLICANT: Nishida, Ken'ichi
; TITLE OF INVENTION: Endothelial Nitric Oxide Synthase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/908,245
; FILING DATE: 19920702
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU 111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4089 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bovine
; TISSUE TYPE: Aorta
; CELL TYPE: Endothelial
US-07-908-245-1

Query Match 76.5%; Score 13; DB 1; Length 4089;
Best Local Similarity 76.5%; Pred. No. 85;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGNA 17
|||:|||||
Db 3840 GACTTAGGTGGGAAA 3824

RESULT 10
US-09-123-708-5/c
; Sequence 5, Application US/09123708
; Patent No. 6146887
; GENERAL INFORMATION:
; APPLICANT: SCHRADER, Juergen
; APPLICANT: GODECKE, Axel
; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN GENE THERAPEUTIC
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
; FILE REFERENCE: 511169-2003
; CURRENT APPLICATION NUMBER: US/09/123,708
; CURRENT FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: 08/553,503

EARLIER FILING DATE: 1996-03-01
EARLIER APPLICATION NUMBER: P4411402.8
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 4097
TYPE: DNA
ORGANISM: Cytomegalovirus
US-09-123-708-5

Query Match 76.5%; Score 13; DB 3; Length 4097;
Best Local Similarity 76.5%; Pred. No. 85;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|:| || ||||| ||

DB 3868 GACTTAGGCTGGGAAA 3852

RESULT 11

US-09-123-624-5/c
Sequence 5, Application US/09123624
Patent No. 6149936
GENERAL INFORMATION:
APPLICANT: SCHRADER, Jurgen
APPLICANT: GODECKE, Axel
TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN THE GENE THERAPEUTIC
TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
FILE REFERENCE: 511169-2004
CURRENT APPLICATION NUMBER: US/09/123,624
CURRENT FILING DATE: 1998-07-28
PRIOR APPLICATION NUMBER: 08/553,503
PRIOR FILING DATE: 1996-03-01
PRIOR APPLICATION NUMBER: 4411402.8
PRIOR FILING DATE: 1994-03-31
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 4097
TYPE: DNA
ORGANISM: Bos taurus
US-09-123-624-5

Query Match 76.5%; Score 13; DB 3; Length 4097;
Best Local Similarity 76.5%; Pred. No. 85;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|:| || ||||| ||

DB 3868 GACTTAGGCTGGGAAA 3852

RESULT 12

US-08-822-028-74
Sequence 74, Application US/08822028
Patent No. 5993813
GENERAL INFORMATION:
APPLICANT: MEZES, PETER S
APPLICANT: GOURLIE, BRIAN B
APPLICANT: RIXON, MARK W
APPLICANT: ANDERSON, WH KERR
APPLICANT: KAPLAN, DONALD A
APPLICANT: SCHOLOM, JEFFREY
TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
MODIFIED ANTIBODIES FOR CANCER TREATMENT
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: DUANE C ULMER
STREET: P.O. BOX 1967
CITY: MIDLAND

STATE: MICHIGAN
COUNTRY: USA
ZIP: 48641-1967
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/822,028
FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/040,687

FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: ULMER, DUANE C
REGISTRATION NUMBER: 34,941

REFERENCE/DOCKET NUMBER: C-37,075C
TELECOMMUNICATION INFORMATION:

TELEPHONE: (517) 636-8104

INFORMATION FOR SEQ ID NO: 74:

SEQUENCE CHARACTERISTICS:

LENGTH: 508 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-822-028-74

Query Match 74.1%; Score 12.6; DB 2; Length 508;

Best Local Similarity 70.6%; Pred. No. 1.1e+02;

Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17

||:|:| || ||||| ||

DB 25 GACTTTGGCTGGGGCA 4.

RESULT 13

US-08-479-285-74
Sequence 74, Application US/08479285
Patent No. 6207815
GENERAL INFORMATION:
APPLICANT: MEZES, PETER S
APPLICANT: GOURLIE, BRIAN B
APPLICANT: RIXON, MARK W
APPLICANT: ANDERSON, WH KERR
APPLICANT: KAPLAN, DONALD A
APPLICANT: SCHOLOM, JEFFREY
TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
MODIFIED ANTIBODIES FOR CANCER TREATMENT
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: DUANE C ULMER
STREET: P.O. BOX 1967
CITY: MIDLAND
STATE: MICHIGAN
COUNTRY: USA
ZIP: 48641-1967
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/479,285
FILING DATE: 07-JUN-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/040687
FILING DATE: 31-MAR-1993
ATTORNEY/AGENT INFORMATION:

NAME: ULMER, DUANE C
 REGISTRATION NUMBER: 34,941
 REFERENCE/DOCKET NUMBER: C-37,075C
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (517) 636-8104
 INFORMATION FOR SEQ ID NO: 74:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 508 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-479-285-74

Query Match 74.1%; Score 12.6; DB 4; Length 508;
 Best Local Similarity 70.6%; Pred. No. 1.1e+02;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGNAA 17
 ||:|||||
 Db 25 GACTTTGGCTGGGGCA 41

RESULT 14
 US-08-599-171A-24
 : Sequence 24, Application US/08599171A
 : Patent No. 5814473
 : GENERAL INFORMATION:
 : APPLICANT: WARREN, Patrick V.
 : TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
 : NUMBER OF SEQUENCES: 32
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: CARELIA, BYRNE, BAIN, GILFILLAN,
 : ADDRESSEE: CECCHI, STEWART & OLSTEIN
 : STREET: 6 BECKER FARM ROAD
 : CITY: ROSELAND
 : STATE: NEW JERSEY
 : COUNTRY: USA
 : ZIP: 07068
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5 INCH DISKETTE
 : COMPUTER: IBM PS/2
 : OPERATING SYSTEM: MS-DOS
 : SOFTWARE: WORD PERFECT 5.1
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/599.171A
 : FILING DATE: Concurrently
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER:
 : FILING DATE:
 : CLASSIFICATION:
 : ATTORNEY/AGENT INFORMATION:
 : NAME: HERRON, CHARLES J.
 : REGISTRATION NUMBER: 28,019
 : REFERENCE/DOCKET NUMBER: 331400-38
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 201-994-1700
 : TELEFAX: 201-994-1744
 : INFORMATION FOR SEQ ID NO: 24:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 912 NUCLEOTIDES
 : TYPE: NUCLEIC ACID
 : STRANDEDNESS: SINGLE
 : TOPOLOGY: LINEAR
 : MOLECULE TYPE: GENOMIC DNA
 : US-08-599-171A-24

Query Match 74.1%; Score 12.6; DB 1; Length 912;
 Best Local Similarity 70.6%; Pred. No. 1.2e+02;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGNAA 17
 ||:|||||
 Db 686 GACTTCGGGTGGAGAA 702

RESULT 15
 US-08-646-590B-24
 : Sequence 24, Application US/08646590B
 : Patent No. 5962283
 : GENERAL INFORMATION:
 : APPLICANT: Warren, Patrick V.
 : APPLICANT: Swanson, Ronald V.
 : TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
 : NUMBER OF SEQUENCES: 42
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Fish & Richardson, P.C.
 : STREET: 4225 Executive Square, Suite 1400
 : CITY: La Jolla
 : STATE: CA
 : COUNTRY: US
 : ZIP: 92037
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: Windows95
 : SOFTWARE: FastSEQ for Windows Version 2.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/646.590B
 : FILING DATE: 08-May-1996
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/599,171
 : FILING DATE: 09-FEB-1996
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: PCT/US97/01094
 : FILING DATE: 21-January-1997
 : CLASSIFICATION: 435
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Haile, Ph.D., Lisa A.
 : REGISTRATION NUMBER: 38,347
 : REFERENCE/DOCKET NUMBER: 09010/017001
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 619/678-5070
 : TELEFAX: 619/678-5099
 : INFORMATION FOR SEQ ID NO: 24:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 912 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: Genomic DNA
 : FEATURE:
 : NAME/KEY: Coding Sequence
 : LOCATION: 1...909
 : US-08-646-590B-24

Query Match 74.1%; Score 12.6; DB 2; Length 912;
 Best Local Similarity 70.6%; Pred. No. 1.2e+02;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGNAA 17
 ||:|||||
 Db 686 GACTTCGGGTGGAGAA 702

Search completed: November 5, 2001, 18:05:11
 Job time: 15111 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:01:04 ; Search time 5816.79 Seconds
(without alignments)
27.627 Million cell updates/sec

Title: US-08-894-356C-22

Perfect score: 17

Sequence: 1 GAYTTYGGNTGGGNA 17

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est33:*
- 26: gb_est34:*
- 27: gb_est35:*
- 28: gb_est36:*
- 29: gb_est37:*
- 30: gb_est38:*
- 31: gb_est39:*
- 32: gb_est40:*
- 33: em_estba:*
- 34: em_estfun:*
- 35: em_esthum1:*
- 36: em_esthum2:*
- 37: em_esthum3:*
- 38: em_esthum4:*
- 39: em_esthum5:*
- 40: em_esthum6:*
- 41: em_esthum7:*
- 42: em_esthum8:*
- 43: em_esthum9:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estrol:*
- 83: em_estrol:*
- 84: em_estrol:*
- 85: em_estrol:*
- 86: em_estrol:*
- 87: em_estrol:*
- 88: em_estrol:*
- 89: em_estrol:*
- 90: em_estrol:*
- 91: em_estrol:*
- 92: em_estrol:*
- 93: em_estrol:*
- 94: em_estrol:*
- 95: em_estrol:*
- 96: em_estrol:*
- 97: em_estrol:*
- 98: em_estrol:*
- 99: em_estrol:*
- 100: em_estrol:*
- 101: em_estrol:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.


```

AW234169      244 bp      mRNA      EST      17-JUL-2000
LOCUS      sf22a08.y1 Gm-cl028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
DEFINITION      Gm-cl028-687 5' similar to TR:Q43583 Q43583 HSR201 PROTEIN. ;, mRNA
sequence.
ACCESSION      AW234169
VERSION
KEYWORDS
SOURCE
ORGANISM      soybean.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
REFERENCE
AUTHORS      1 (bases 1 to 244)
Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvett,V., Khanna
,A., Bolla,B., Marr,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wyllie,T., Underwood,K., Steptoe,W., Theising,B., Allen,M., Bowers
,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Insert Length: 486 Std Error: 0.00.
FEATURES
source
1. .244
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl028-687"
/clone_lib="Gm-cl028"
/tissue_type="roots of 'Supernod' plants"
/lab_host="DH10B"
/notes="vector: pBluescript II XR; Site_1: EcoRI; Site_2:
XhoI; The mRNA was isolated from roots of Glycine max
'Supernod' plants generously donated by Dr. Gary Stacey.
The seedlings were inoculated with Bradyrhizobium
japonicus, strain USDA110 prior to harvest. Stratagene's
cDNA synthesis kit (catalog number 200401) was used to
synthesize the cDNA. First-strand synthesis was performed
with 5-methyl dCTP, hence the ligated cDNA was
hemimethylated. A modification of Stratagene's
first-strand synthesis primer was used. An 'anchor'
nucleotide (V-A-C, or G) was added to the 3' end of the
primer (GAGAGAGAGAGAGAGAGACTAGTCGAG(T)18V) to anchor
the primer at the 5' end of the poly(A) tract. After
second-strand synthesis, the cDNA ends were filled in with
cloned Pfu DNA polymerase, ligated to EcoRI adapters and
subsequently phosphorylated. The XhoI site within the
first-strand synthesis primer was then restricted by
digestion with XhoI; all XhoI sites in the cDNA would be
protected by their hemimethylated status. The cDNA
constructs were size-fractionated with a 500bp cutoff,
using GibcoBRL Life Technologies' cDNA size fractionation
column. The column eluent was then ligated into
Stratagene's pBluescript II XR Predigested vector
(pBluescript II SK(+)) that has been digested with EcoRI
and XhoI, and phosphorylated by Stratagene). Both the
white and blue colonies appear to contain recombinant
plasmids with cDNA inserts, based on size (n=25). This
library was constructed by Dr. Paul Keim and Dr. Virginia
Corvett."
65 a 37 c 65 g 77 t

BASE COUNT
Query Match 83.5%; Score 14.2; DB 21; Length 247;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GAYTTYGGTGGGNA 17
||:|:|:|:|:|:|:|:|:|
Db 10 GACITTTGGTTGGGGAA 26
||:|:|:|:|:|:|:|:|:|

RESULT 3
LOCUS      AI488429/c
DEFINITION      EST245768 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
CLSD21D13, mRNA sequence.
ACCESSION      AI488429
VERSION
KEYWORDS
SOURCE      tomato.
ORGANISM      Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE
AUTHORS      1 (bases 1 to 247)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
,Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
,S.D. and Giovannoni,J.
Generation of ESTs from tomato carpel tissue
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU.
Location/Qualifiers
1. .247
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLSD21D13"
/clone_lib="tomato ovary, TAMU"
/tissue_type="carpel"
/dev_stage="5 days pre-anthesis to 5 days post-anthesis"
/lab_host="XLI-Blue MRF"
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLED - Tomato Carpel EST Library_ OligodT-primed and
directionally cloned cDNA in vector Lambda ZAP II with 5'
and 3' ends located at the EcoRI and XhoI sites,
respectively."
72 a 67 c 27 g 81 t

BASE COUNT
Query Match 83.5%; Score 14.2; DB 21; Length 247;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GAYTTYGGTGGGNA 17
||:|:|:|:|:|:|:|:|:|
Db 222 GATTITGGATGGGAA 206
||:|:|:|:|:|:|:|:|:|

RESULT 4
LOCUS      BG550470
DEFINITION      BG550470 247 bp mRNA EST 09-APR-2001
947076G02.x2 947 - 2 week shoot from Barkan lab zea mays cDNA, mRNA
sequence.

```

```

ACCESSION BG550470
VERSION BG550470.1 GI:13562250
KEYWORDS EST.
SOURCE Zea mays.
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 247)
AUTHORS Walbot V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 947076 row: G column: 02.
FEATURES
source Location/Qualifiers
1..247
/organism="Zea mays"
/cultivar="B73"
/db_xref="taxon:4577"
/clone_lib="947 - 2 week shoot from Barkan lab"
/tissue_type="leaf and stem, including leaf base"
/dev_stage="2 week old seedling (3 leaves)"
/lab_host="X11-Blue"
/note="Organ: shoot; Vector: Lambda ZAP (pBlueScript SK-);
Site:1: EcoRI; Site:2: XhoI; Directionally cloned using
Stratagene's Unizap XR cDNA cloning kit with the 5' end
at the EcoRI site. The library represents 8 x 105
independent recombinant phage. The plants were greenhouse
grown."
BASE COUNT 57 a 63 c 85 g 42 t
ORIGIN
Query Match 83.5%; Score 14.2; DB 155; Length 247;
Best Local Similarity 76.5%; Pred. NO. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 GAYTTGGNTGGGNA 17
||:|||||
Db 224 GATTTGGTGGGAAA 240

RESULT 5
AZ015624 249 bp DNA GSS 25-FEB-2000
LOCUS RPI-23-258D7.TJ RPI-23 Mus musculus genomic clone RPI-23-258D7,
DEFINITION DNA sequence.
ACCESSION AZ015624
VERSION AZ015624.1 GI:7091008
KEYWORDS house musculus.
SOURCE Mus musculus.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 249)
AUTHORS Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Aklnret
,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
and Fraser,C.M.
TITLE Mouse BAC End Sequences from Library RPI-23
JOURNAL Unpublished (1999)
COMMENT Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200

```

```

Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 258 row: D column: 7
Seq primer: SP6
Class: BAC ends.
FEATURES
source Location/Qualifiers
1..249
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="RPI-23-258D7"
/clone_lib="RPI-23"
/sex="Female"
/lab_host="DH10B"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site:1:
EcoRI; Site:2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methylase. Size
selected DNA was cloned into the pBACe3.6 vector at the
EcoRI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT 87 a 54 c 39 g 69 t
ORIGIN
Query Match 83.3%; Score 14.2; DB 237; Length 249;
Best Local Similarity 76.3%; Pred. NO. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 GAYTTGGNTGGGNA 17
||:|||||
Db 60 GATTTGGTGGGCAA 76

RESULT 6
AA650754 258 bp mRNA EST 31-OCT-1997
LOCUS 30868 Lambda-PRL2 Arabidopsis thaliana cDNA clone 281H6T7, mRNA
DEFINITION sequence.
ACCESSION AA650754
VERSION AA650754.1 GI:2580846
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 258)
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
,E. and Somerville,C.
TITLE Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
JOURNAL Plant Physiol. 106. 1241-1255 (1994)
MEDLINE 95148729
COMMENT Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@lbn.cl.msu.edu
Seq primer: T7 dye primer.
FEATURES
source Location/Qualifiers
1..258
/organism="Arabidopsis thaliana"

```

/strain="var columbia"
 /db_xref="taxon:3702"
 /clone="281H67"
 /clone_lib="Lambda-PRL2"
 /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
 Lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques. The vector is BRL's lambda Zip-Lox. The cDNA
 inserts were directionally cloned with Sal-Not arms using
 oligo dt primed cDNA. "

BASE COUNT 74 a 42 c 62 g 70 t 10 others
 ORIGIN

Query Match 83.5%; Score 14.2; DB 10; Length 258;
 Best Local Similarity 76.5%; Pred. No. 1.2e+03;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GAYTTGGTGGGNA 17
 ||:||||| ||||| ||
 Db 53 GATTTGGTGGGTAA 69
 ||:||||| ||||| ||

RESULT 7
 BG046171
 LOCUS
 DEFINITION BG046171 259 bp mRNA EST 25-JAN-2001
 ID: Gm-cl060-215 5' similar to TR:Q43583 Q43583 HSR201 PROTEIN. ;,
 mRNA sequence.

ACCESSION BG046171
 VERSION BG046171.1 GI:12494655
 KEYWORDS EST.
 SOURCE soybean.
 ORGANISM Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
 1 (bases 1 to 259)
 Shoemaker,R., Kelm,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
 ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
 ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
 ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
 ,R., Waterston,R. and Wilson,R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134 For further information
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or
 info@genomesystems.com web site: www.genomesystems.com
 High quality sequence stop: 157.

FEATURES
 source
 1. 259
 Location/Qualifiers
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl060-215"
 /clone_lib="Gm-cl060"
 /tissue_type="Root, 2 week seedlings"
 /lab_host="DH10B"
 /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
 XhoI; The cDNA library was constructed from mRNA isolated

from root tissue of 2 week old seedlings for PI468916.
 Complementary DNA was synthesized from mRNA using a primer
 consisting of a poly(dT) sequence with a XhoI restriction
 site. EcoRI adapters were ligated to the blunt-ended cDNA
 fragments followed by XhoI digestion. The cDNA fragments
 were directionally cloned into the EcoRI-XhoI restriction
 site of the pBluescript vector. The ligated cDNA fragments
 were transformed into DH10B host cells (Gibco BRL). This
 library was constructed in the laboratory of Dr. Randy
 Shoemaker at Iowa state university."

BASE COUNT 50 a 41 c 91 g 77 t
 ORIGIN

Query Match 83.5%; Score 14.2; DB 173; Length 259;
 Best Local Similarity 76.5%; Pred. No. 1.2e+03;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GAYTTGGTGGGNA 17
 ||:||||| ||||| ||
 Db 89 GACTTTGGTGGGGAA 105
 ||:||||| ||||| ||

RESULT 8
 AA651035
 LOCUS
 DEFINITION AA651035 263 bp mRNA EST 05-JAN-1998
 31149 Lambda-PRL2 Arabidopsis thaliana cDNA clone 168A5XP 3', mRNA
 sequence.

ACCESSION AA651035
 VERSION AA651035.1 GI:2581127
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 263)
 Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
 ,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel
 ,E. and Somerville,C.
 Genes galore: a summary of methods for accessing results from
 large-scale partial sequencing of anonymous Arabidopsis cDNA clones
 Plant Physiol. 106, 1241-1255 (1994)
 95148729

CONTACT: Thomas Newman
 MSD-DOE Plant Research Laboratory
 Michigan State University
 MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
 Lansing, MI
 Tel: 517-353-0854
 Fax: 517-353-9168
 Email: 22313tcn@bm.cl.msu.edu
 The sequence entry for this EST has been reverse complemented and
 is being submitted in the sense orientation.
 Seq primer: M13-Universal.

FEATURES
 source
 1. 263
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /strain="var columbia"
 /db_xref="taxon:3702"
 /clone="168A5XP"
 /clone_lib="Lambda-PRL2"
 /note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not;
 Lambda PRL2 is a cDNA library derived from equal
 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
 day germinated etiolated seedlings; 2) tissue culture
 grown roots; 3) staged plants half with 24 hour light
 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
 same plants as 3 but aerial tissue (stems, flowers and
 siliques. The vector is BRL's lambda Zip-Lox. The cDNA
 inserts were directionally cloned with Sal-Not arms using
 oligo dt primed cDNA. "

BASE COUNT 56 a 68 c 50 g 79 t 10 others

ORIGIN

Query Match 83.5%; Score 14.2; DB 10; Length 263;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
II:II:II IIIII II

Db 90 GATTTCGGATGGGCAA 74

RESULT 9
AI608251/c

LOCUS
DEFINITION
v93f09 x1 Knowles Solter mouse embryonic stem cell Mus musculus
21-APR-1999
EST

AI608251
CDNA clone IMAGE:894569 3', mRNA sequence.

AI608251
AI608251.1 GI:4617418

EST.

SOURCE
house mouse.

ORGANISM
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 278)
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person
B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter
E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R. and Wilson, R.
The WashU-NCI Mouse EST Project 1999

Unpublished (1999)

CONTACT: Marra M/WashU-NCI Mouse EST Project 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:522529

This clone was previously sequenced on the 5' end only, this new
data is from the 3' end

Seq primer: Primer name ambiguous

High quality sequence stop: 274.

Location/Qualifiers

1..278

/organism="Mus musculus"

/strain="B6D2 F1/J"

/db_xref="taxon:10090"

/clone="IMAGE:894569"

/clone_lib="Knowles Solter mouse embryonic stem cell"

/dev_stage="embryo"

/lab_host="DH10B"

/note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; Cloned
unidirectionally from mRNA prepared from 800 blastocysts.
Primer: SalI(dT): 5'-CGGTCGACCGTCGACCGTTTTTTTTTTT-3'.

CDNAs were cloned into the NotI/SalI sites of a pSPORT
vector (Life Technologies).

57 a 75 c 48 g 98 t

BASE COUNT 57 a 75 c 48 g 98 t

ORIGIN

Query Match 83.5%; Score 14.2; DB 22; Length 278;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
II:II:II IIIII II

Db 99 GATTTTGGTGGGCAA 83

RESULT 10

BB264508/c

LOCUS

DEFINITION

musculus CDNA clone A830015G22 3', mRNA sequence.

ACCESSION

BB264508

VERSION

BB264508.1

KEYWORDS

EST.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

AUTHORS

1 (bases 1 to 294)

Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurlhar, C., Okakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tominaga, N., Toya, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, K., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Konno, H., et al.)

Unpublished (2000)

CONTACT: Yoshihide Hayashizaki

Genome Exploration Research Group, Life Science Tsukuba Center,

Genome Science Laboratory

The Institute of Physical and Chemical Research (RIKEN), Genomic

Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: +81-298-36-9013

Fax: +81-298-36-9098

Email: genome-res@rtc.riken.go.jp,

URL: http://genome.rtc.riken.go.jp/

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Thermotabilization and thermoactivation of thermolabile enzymes by

trehalose and its application for the synthesis of full length

CDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Itoh, M., Kitsumai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

19-44 (1999)

Please visit our web site (http://genome.rtc.riken.go.jp) for

further details.

FEATURES

source

1..294

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="A830015G22"

/clone_lib="RIKEN full-length enriched, 10 days neonate

cortex"

/tissue_type="cortex"

/dev_stage="10 days neonate"

/lab_host="DH10B"

/note="Site_1: SalI; Site_2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in

RIKEN. Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5'

GAGAGAGACGATCCAGACCTCTTTTTTTTTTTTTT-3'], cDNA was

prepared by using trehalose thermo-activated reverse

transcriptase and subsequently enriched for full-length by

cap-trapper. cDNA went through one round of normalization

to Rot = 20.0 and subtraction to Rot = 459.0. Second

BB264508 294 bp mRNA

musculus CDNA clone A830015G22 3', mRNA sequence.

BB264508

BB264508.1 GI:8960965

EST.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 294)

Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurlhar, C., Okakabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tominaga, N., Toya, T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I., Yano, K., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Konno, H., et al.)

Unpublished (2000)

CONTACT: Yoshihide Hayashizaki

Genome Exploration Research Group, Life Science Tsukuba Center,

Genome Science Laboratory

The Institute of Physical and Chemical Research (RIKEN), Genomic

Sciences Center

3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: +81-298-36-9013

Fax: +81-298-36-9098

Email: genome-res@rtc.riken.go.jp,

URL: http://genome.rtc.riken.go.jp/

Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoka, S., Sasaki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Thermotabilization and thermoactivation of thermolabile enzymes by

trehalose and its application for the synthesis of full length

CDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)

Itoh, M., Kitsumai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y. and Hayashizaki, Y.

Automated filtration-based high-throughput plasmid preparation

system. Genome Res. 9 (5), 463-470 (1999)

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,

19-44 (1999)

Please visit our web site (http://genome.rtc.riken.go.jp) for

further details.

Location/Qualifiers

1..294

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="A830015G22"

/clone_lib="RIKEN full-length enriched, 10 days neonate

cortex"

/tissue_type="cortex"

/dev_stage="10 days neonate"

/lab_host="DH10B"

/note="Site_1: SalI; Site_2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in

RIKEN. Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5'

GAGAGAGACGATCCAGACCTCTTTTTTTTTTTTTT-3'], cDNA was

prepared by using trehalose thermo-activated reverse

transcriptase and subsequently enriched for full-length by

cap-trapper. cDNA went through one round of normalization

to Rot = 20.0 and subtraction to Rot = 459.0. Second

strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTTAATAATTAATCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 64 a 77 c 66 g 87 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 129; Length 294;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTGGTGGGGNAA 17
||:|:|:|:|:|:|:|:|:|
Db 243 GACTTTGGTGGGGAA 227

RESULT 11
AM471629
LOCUS AM471629 307 bp mRNA EST 24-FEB-2000
DEFINITION sll3h01.v1 Gm-cl029 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl029-1130 5' similar to TR:004201 004201
HYPERSENSITIVITY-RELATED GENE 201 ISOLOG. ;, mRNA sequence.

ACCESSION AM471629
VERSION AM471629.1 GI:7041735
KEYWORDS EST.
SOURCE soybean.

ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE 1 (bases 1 to 307)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpeliding,J., Corvett,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)

TITLE Public Soybean EST Project
JOURNAL
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
High quality sequence stop: 301.

FEATURES
source
1..307
Location/Qualifiers
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl029-1130"
/clone_lib="Gm-cl029"
/tissue_type="very young cotyledons of greenhouse grown plants"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: Not I; Site_2: Sal I; This cDNA library was constructed from mRNA isolated from very young cotyledons (20-50mgs fresh weight) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction

site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed by Dr. Paul Keim and Dr. Virginia Corvett."

BASE COUNT 96 a 33 c 96 g 82 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 116; Length 307;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTGGTGGGGNAA 17
||:|:|:|:|:|:|:|:|:|
Db 127 GATTTGGATGGGGAA 143

RESULT 12
AI856623
LOCUS sb39g12.v1 Gm-cl014 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl014-95 5' similar to TR:004201 004201
HYPERSENSITIVITY-RELATED GENE 201 ISOLOG. ;, mRNA sequence.

ACCESSION AI856623
VERSION AI856623.1 GI:5510239
KEYWORDS EST.
SOURCE soybean.

ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE 1 (bases 1 to 324)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpeliding,J., Corvett,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)

TITLE Public Soybean EST Project
JOURNAL
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
Seq primer: -40RP from Gibco
High quality sequence stop: 311.

FEATURES
source
1..324
Location/Qualifiers
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl014-95"
/clone_lib="Gm-cl014"
/tissue_type="Leaves, 2-3 week old seedlings, greenhouse grown"
/lab_host="DH10B"
/note="Vector: pT73Pac (pT73, Pharmacia); Site_1: EcoRI; Site_2: HindIII; This cDNA library was constructed from mRNA isolated from leaves of 2-3 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a 3' anchored poly (dT) primer. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and HindIII. The cDNA fragments were directionally cloned into the EcoRI-HindIII restriction site of the pT73-Pac vector. The ligated cDNA fragments

were transformed into DH10B host cells (Gibco BRL). This library was constructed by Dr. Randy Shoemaker and Dr. John Erpellding."

BASE COUNT 93 a 49 c 101 g 79 t 2 others
ORIGIN

Query Match 83.5%; Score 14.2; DB 102; Length 324;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNAA 17
||:||:|||||
Db 192 GATTTGGGTGGGGAA 208

RESULT 13
AZ813888 331 bp DNA GSS 20-FEB-2001
LOCUS 2M0081B21F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC2M0081B21 F, DNA sequence.
ACCESSION AZ813888
VERSION AZ813888.1 GI:12983784
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 331)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D. Wells R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0081 row: B column: 21
Seq primer: CGTTGTAACGACGCCAGT
Class: plasmid ends
High quality sequence stop: 331.

FEATURES
source
1. .331
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0081B21"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="F. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: pW42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrotynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pW42 (g114732141gb1Ar129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 75 a 82 c 78 g 96 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 250; Length 331;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNAA 17
||:||:|||||
Db 115 GATTTGGGTGGGGAA 131

RESULT 14
AQ034368/C 333 bp DNA GSS 11-JUL-1998
LOCUS CIT-HSP-2319K4.TF CIT-HSP Homo sapiens genomic clone 2319K4, DNA
DEFINITION sequence.
ACCESSION AQ034368
VERSION AQ034368.1 GI:3300542
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 333)
AUTHORS Adams,K., Berry,K., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Golden,K., Berry,K., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamads@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.

FEATURES
source
1. .333
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2319K4"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelOBAC11; Site_1: HindIII; Site_2: HindIII"
BASE COUNT 90 a 100 c 55 g 88 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 223; Length 333;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNAA 17
||:||:|||||
Db 237 GATTTGGGTGGGGAA 221

RESULT 15
BE460581/c

LOCUS BE460581 347 bp mRNA EST 27-JUL-2000
DEFINITION EST412000 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG31N6, mRNA sequence.
ACCESSION BE460581
VERSION BE460581.1 GI:9504883
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 347)
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
S.D.
TITLE Generation of ESTs from tomato fruit tissue, breaker stage
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu
5 prime sequence.
FEATURES
Source
1..347
Location/Qualifiers
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG31N6"
/clone_lib="tomato breaker fruit, TIGR"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmCvadapt; Site.1: EcoRI;
Site.2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."
BASE COUNT 120 a 71 c 40 g 116 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 167; Length 347;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GAYTTYGGNTGGGGNAA 17
||:|||||
Db 91 GATTTTGGATGGGGAAA 75
||:|||||

Search completed: November 5, 2001, 18:01:08
Job time: 14908 sec